

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 27, 2003, 21:11:47 / Search time 8407 Seconds

(without alignments)
1942.033 Million cell updates/sec

Title: US-09-821-687-4

Perfect score: 2982

Sequence: 1 MATHEVNGTEEPMDTSTA.....AQQGRGGKGYEAGPDLLQ 561

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgcn2.1/USPTO.spool/US09821687/runat.22042003.165241.7951/app.query.fasta_1.711
-DB=GenEmbl -QFMT=fastap -SUFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -List=45
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PCT -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821687@CGCN 1.1.5173 -runat.22042003.165241.7951 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_da:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_ov:.*
7: gb_ov:.*
8: gb_ov:.*
9: gb_ov:.*
10: gb_ov:.*
11: gb_ov:.*
12: gb_ov:.*
13: gb_ov:.*
14: gb_ov:.*
15: gb_ov:.*
16: gb_ov:.*
17: gb_ov:.*
18: gb_ov:.*
19: gb_ov:.*
20: gb_ov:.*
21: gb_ov:.*
22: gb_ov:.*
23: gb_ov:.*
24: gb_ov:.*
25: gb_ov:.*
26: gb_ov:.*
27: gb_ov:.*
28: gb_ov:.*

29: em_v1:.*
30: em_hcg_hum:.*
31: em_hcg_inv:.*
32: em_hcg_other:.*
33: em_hcg_mus:.*
34: em_hcg_pln:.*
35: em_hcg_rod:.*
36: em_hcg_mam:.*
37: em_hcg_vrt:.*
38: em_hcg_hum:.*
39: em_hcg_mus:.*
40: em_hcg_mus:.*
41: em_hcg_other:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2982	100.0	3452	10 AB035725	AB035725 Mus muscu
2	2964.5	99.4	2221	9 AF155568	AF155568 Homo sapi
3	2964	99.4	1686	9 AY034483	AY034483 Homo sapi
4	2922	98.0	1872	9 AF034481	AF034481 Homo sapi
5	2922	98.0	2932	9 AF037448	AF037448 Homo sapi
6	2883	96.7	2373	10 AF093821	AF093821 Mus muscu
7	2882	96.6	2079	6 AR177851	AR177851 Sequence
8	2752	92.3	3319	9 BC032643	BC032643 Homo sapi
9	2705.5	90.7	1767	9 AY034482	AY034482 Homo sapi
10	2676	88.7	103819	9 HS0775C13	AL109618 Human DNA
11	2676	89.7	212659	2 AL390737	AL390737 Homo sapi
12	2604	87.3	2232	10 AF408434	AF408434 Mus muscu
13	2502	83.9	3101	6 AX011753	AX011753 Sequence
14	2413.5	80.9	1899	10 AF441128	AF441128 Mus muscu
15	2403	80.6	2644	9 AF000364	AF000364 Homo sapi
16	2391.5	80.2	2722	9 BC001449	BC001449 Homo sapi
17	2275	78.3	2471	10 BC004679	BC004679 Mus muscu
18	2114	70.9	181810	2 AC126148	AC126148 Rattus no
19	2066.5	69.3	121304	2 AC079134	AC079134 Mus muscu
20	2066.5	69.3	136488	2 AC084410	AC084410 Mus muscu
21	1745	58.5	139996	9 AL161799	AL161799 Human DNA
22	1520	51.0	221309	2 AC122217	AC122217 Mus muscu
23	1468	49.2	151297	2 AC118904	AC118904 Rattus no
24	1382	46.3	934	6 AX013715	AX013715 Sequence
25	1360	45.6	129237	2 AC127932	AC127932 Rattus no
26	1301	43.6	2606	3 AY058477	AY058477 Drosophila
27	1175	39.4	202374	10 AC122868	AC122868 Mus muscu
28	1172.5	39.3	110783	2 AC129693	AC129693 Rattus no
29	1159.5	38.9	118429	2 AC119124	AC119124 Rattus no
30	1041	34.9	165311	2 AC112849	AC112849 Rattus no
31	964	32.3	2473	9 AF262323	AF262323 Homo sapi
32	963.5	32.3	2498	10 BC013481	BC013481 Mus muscu
33	959	32.2	2218	4 CFA388555	AJ388555 Canis fam
34	955.5	32.0	2585	10 BC034195	BC034195 Mus muscu
35	908.5	30.5	4020	9 AK000280	AK000280 Homo sapi
36	903	30.3	159699	2 AC026833	AC026833 Homo sapi
37	902.5	30.3	188295	9 AC098869	AC098869 Homo sapi
38	902.5	30.3	248201	2 AC115293	AC115293 Mus muscu
39	896	30.0	1863	10 AF290984	AF290984 Rattus no
40	894	30.0	1785	10 AY028945	AY028945 Rattus no
41	883.5	29.6	1794	9 HSA27078	AJ27078 Homo sapi
42	881.5	29.6	1945	9 AF209192	AF209192 Homo sapi
43	879.5	29.5	1770	9 HSA27079	AJ27079 Homo sapi
44	876.5	29.4	1761	9 AF271789	AF271789 Homo sapi
45	876.5	29.4	128739	2 AC116602	AC116602 Mus muscu

ALIGNMENTS

RESULT 1

AB035725
LOCUS AB035725 3452 bp mRNA linear ROD 31-MAR-2000
DEFINITION Mus musculus SYNCRIP mRNA, complete cds.
ACCESSION AB035725
VERSION AB035725.1 GI:6576814
KEYWORDS SYNCRIP.
SOURCE Mus musculus (strain:DDY) postnatal 21 day cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Mizutani, A., Fukuda, M., Ibata, K., Shiraiishi, Y., and Mikoshiba, K.
TITLE SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear
ribonucleoprotein R, interacts with ubiquitous synaptoctagmin
isoforms
JOURNAL J. Biol. Chem. 275 (13), 9823-9831 (2000)
MEDLINE 20200483
REFERENCE 2 (bases 1 to 3452)
AUTHORS Mizutani, A., Fukuda, M. and Mikoshiba, K.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1999) Akihiro Mizutani, Brain Science Institute,
RIKEN, the Laboratory for Developmental Neurobiology, 2-1 Hirosawa,
Wako, Saitama 351-0198, Japan (E-mail: amizutani@ims.u-tokyo.ac.jp),
Tel: +81-48-467-9745, Fax: +81-48-467-9744)
FEATURES
source location/Qualifiers
1..3452
/organism="Mus musculus"
/strain="DDY"
/db_xref="taxon:10090"
/dev_stage="postnatal 21 day"
1..3452
/gene="SYNCRIP"
1..3452
/gene="SYNCRIP"
/evidence=experimental
154..1839
/gene="SYNCRIP"
/protein_start=1
/protein_id="BA088342.1"
/db_xref="GI:6576815"
translation="MATEHVNGCTEPMDDTSAAVHSENFQTLDAGLPQVKAETLD
EIVAGLVNASHDDEBIAIBALKEFNEDGALAVLQFPDSDLSHVOKSAFLCGWMTY
ROBKQGTVAADSSGKDEAKILALERTYTLDTVGKRGKGGPPDSVYSGQPSV
GTEIFVKGKIPRDLFEBELVPLFKAGVIMPLRLMDPLTGANKGVAFTVPTCEAAOS
AVLVNHHETRSGKHIGVCI SVANNPLFVCSIPKSTKQILIEBPGKVBGLTDVILY
HOPDDKKRGRGPFLEIEBHKTAAQARRRLMSGKVKYGVNGVTEWADPIEDDEPM
AKQKVLVFNRLAVTVEELIEKSPSQGKLERVKLKVAFHFDERDGVKAMEBMN
GKDLBENIEIVPAKPPDQKREKKAQRQAKNQMDYVYGPPIHPPPTRRGRGG
RGGYGYPDYGYEDYGYDYNHNRGGYEDPYGYEDFOVGARGRGGRGARGAAP
SRGRGAAPRGARGYSORSGRGSARGVARGAGAAQQRGRGKGVARGPDLQ"

BASE COUNT 1056 a 623 c 777 g 996 t
ORIGIN

Alignment Scores:
Pred. No.: 3,36e-131 Length: 3452
Score: 2982.00 Matches: 561
Percent Similarity: 100.00% Conservative: 0
Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-821-687-4 (1-561) x AB035725 (1-3452)

QY 1 MetAAThrGluHisValAsnGlyAsnGlyThrGluGluProMetAspThrThrsSerAla 20
DB 154 ATGGCTACAGAAACATGTTAAAGAAATGGTACTGAAAGCCCACTGATCACTTCCGCA 213
QY 21 ValLeuHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeuProGlnLeuValAla 40
DB 214 GTTATTCATTCAGAAATTTTCAAGCATTCCTTGAATGCTGTTTTCACCAAGAAAGTGGCT 273
QY 11 GlnuLeuAspGluLeuValValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60

DB 274 GAAAACTAGATGAATTTTACGTTCAGAGGCTTACTTGCACATAGTATTTAGATGAAGA 333
QY 61 AlAlIleGluAlaLeuYsgluPheAsnGluAspGlyAlaLeuAlaValLeuGlnGlnPhe 80
DB 334 GCTATCAGAACTTTAAAGAGTTCAATGAAAGCGGCGCATTTGGCAGTCTTCAACGTTT 393
QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuYsglyValMetLys 100
DB 394 AAGACAGATATCTCTCATGTCAGAACAAAGAGCTTTTATGTGAGTCATGAG 453
QY 101 ThrTyrArgGlnArgGluYsglnGlyThrLysValAlaAspSerSerYsglyProAsp 120
DB 454 ACTTACAGGAGAGAGAAAGAAAGAGGAGCAAAAGTACAGACTCTGTAAGAGCAGAT 513
QY 121 GlnAlaLysLysLysAlaLeuLeuGlnGluArgGlyTyrThrLeuAspValThrThrGly 140
DB 514 GAGCAAGATTAAGGACCTTTTGAAAGAAAGACGGCTACACCTTATGATGATCAACGGT 573
QY 141 GlnArgLysTyrGlyValProProAspSerValTyrSerGlyValGlnProSerVal 160
DB 574 CAGAGAGATATGAGAGACCACTCCAGATTCGTTATTCAGGTACAGACCTTCTGTT 633
QY 161 GlyThrGluLysPheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
DB 634 GGCACTGAGATATTTGTGGGAAAGATCCCAAGAGATCTGTTGAGGATGAGCTTTTCCA 693
QY 181 LeuPheGluLysAlaGlyProIleTPAspLeuArgLeuMetLysProLeuThrGly 200
DB 694 TTTATTTGAGAAAGCTGACCTATATGATGATCTTCTGTTATGATGATCGCTCAGTGT 753
QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGluAlaAlaGlnGluAlaVal 220
DB 754 CTCAACAGAGTTATGAGTGTGCTTGTCACTTTTGTACAAAGAACACACAGAGGCTGTT 813
QY 221 LysLeuTyrAsnAsnHisGluLysIleArgSerGlyLysHisIleGlyValCysIleSerVal 240
DB 814 AATCTGATATATATCATGAATATTCGTTCCGGCAACACATGCTGTCATCTCAGTT 873
QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGluGlnIleLeu 260
DB 874 GCCAACAATATGAGCTTTTGTGGGCTCGATTTCTTAAAGATTAACAGAGGACGATTCCT 933
QY 261 GlnGluPheSerLysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnProAsp 280
DB 934 GAGGAATTTGCAAGATGACAGAGGCTTCACAGATCTATTTTATACCAACCACTGAT 993
QY 281 AspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAspHisLysThrAlaAla 300
DB 994 GACAAAGAAAAAACAAGAGCTTTTCTTCTTGAATATGAGATCAACAACAGCTGCC 1053
QY 301 GlnAlaArgArgArgLeuMetSerGlyLysValLysValITPGLysAsnValGlyThrVal 320
DB 1054 CAGGCAAGACGTAGGCTAATGAGTGTAAAGTCAAGTCTGGGGAATGTGGAACTGTT 1113
QY 321 GlnTyrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysValLeu 340
DB 1114 GAGTGGCTGATCTTATTGAAGTCTGATCTGAACTTATGCAAGATTAAGTAAAGTGC 1173
QY 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSerGln 360
DB 1174 TTTGTACGCAACCTTGCACAAACGTTACAGAAAGAAATTTTGAAGAAAGCATTTATCTCAG 1233
QY 361 PheGlyLysLeuGlnGluArgValLysLysLeuLysAspTyrAlaPheIleHisPheAspGlu 380
DB 1234 TTTGGGAAATCGAAGAGAGTAAAGAGCTAAAGATTAATGCTTTCATTCATTTTGTATGAG 1293
QY 381 ArgAspGlyValAlaValLysAlaMetGluGluMetAsnGlyLysAspLeuGluGluAsn 400
DB 1294 AGAGATGGTGTCTCAAGGCTATGAGAAAGATTAAGTGAAGACTTGGAGGAGAAAT 1353
QY 401 IleGluIleValPheAlaLysProProAspGlnLysArgGlyLysArgLysAlaArg 420
DB 1354 ATTGAATTTGTTTCTTGAAGCAACAGATCAAGAAAGAAAGAAAGAAAGCTCAGAG 1413

QY 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrGlyProProHisMetPro 440
 DB 1414 CAGGAGCAAGAAATCAAAAGTATGATGATTAATCAATTAATGATCACTCAATAGCCT 1473
 QY 441 ProProThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 460
 DB 1474 CCCCACCAAGAGAGGTGAGGGGCGTGAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1533
 QY 461 TyrGlyTyrGlyAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGly 480
 DB 1534 TATGATACCAAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1593
 QY 481 TyrGlyAspProTyrTyrGlyTyrGlyAspPheGlnValGlyValAlaArgGlyArgGly 500
 DB 1594 TATGAGATCCATCACTATGTTATGATTAATGATTAATGATTAATGATTAATGATTA 1653
 QY 501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGlyArg 520
 DB 1654 AGAGGAGCAAGGGGCTGCTGCTCATCAAGAGTCTGAGGGCTGCTGCTGCTGCTGAG 1713
 QY 521 AlaGlyTyrSerGlnArgGlyArgGlyProGlySerAlaArgGlyValAlaArgGlyArg 540
 DB 1714 GCGGTTATTCACAGAGAGAGGCGCTGATCAAGAGGCGCTGCGGTGCGAGAGGA 1773
 QY 541 GlyAlaGlnGlnArgGlyArgGlyArgGlyValGlnAlaGlyProAspLeu 560
 DB 1774 GGTGCCCAACAAAGAGGCCGCGGGGGAAGGGGTCAAGGCCGCTGCTGCTGTTA 1833
 QY 561 Gln 561
 DB 1834 CAA 1836
 RESULT 2
 AFI55568 2221 bp mRNA linear PRI 10-JUN-1999
 LOCUS AFI55568 Homo sapiens NSAP1 protein (NSAP1) mRNA, complete cds.
 DEFINITION AFI55568
 ACCESSION AFI55568
 VERSION AFI55568.1 GI:5031511
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Harris, C.E., Boden, R.A. and Astell, C.R.
 TITLE A novel heterogeneous nuclear ribonucleoprotein-like protein
 interacts with NS1 of the minute virus of mice
 J. Virol. 73 (1), 72-80 (1999)
 JOURNAL
 MEDLINE 99102562
 PUBMED 9847309
 REFERENCE
 AUTHORS Harris, C.E., Boden, R.A. and Astell, C.R.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUN-1999) Biochemistry and Molecular Biology, UBC,
 2146 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada
 FEATURES
 source
 1..2221
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="cell line"
 1..2221
 /gene="NSAP1"
 205..1893
 /gene="NSAP1"
 /note="similar to Homo sapiens hn RNP R; probable RNA
 binding protein; interacts with NS1 of the minute virus of
 mice"
 /codon_start=1
 /product="NSAP1 protein"
 /protein_id="AAD38198.1"
 /db_xref="GI:5031512"
 /translation="MATEHVNGNGTEBPMDTTSVAVHSNFQTLIDAGLPQVAKELD

EIVAGLVASHDLDERALTEALKEFNEDGALAVLOFQKSDLSHVNKSAFLGVWKT
 RORRCKGVKVDSSKGPDEAKIKLLEETGTLTUDTQGRYGGAPPPSYSGOOPV
 GTEIFVGKIPDDLEDEDEIVPEKAGPIMDLRLMDPITGLRGAFTPTCTKAQO
 AVKIKYNNHEIRSGHIGVCSVANNRIPVGSIPSKTKEQILIEPSSQVTEGLTVILY
 HQPDDKKKORPCELEVEDHNTAQAORRRRLMSGKVKWGNATVEMADPIEDPPEVA
 AKIVYLFEVRLNANTVTEELIEKASQSKLEKRVKLLKQYALIHDERGAVKASENN
 KQOLEGENIEIVPAKPPDOKERKERRAORQAQNNQMDYDYYGPPHPMPRRGRG
 RGGYVPPDYGVEDYDYGVDYHNTYRGGVEDPYGVEDPYGVEDPYGVEDPYGVEDPYGV
 SRGGAAPPRRAGTISQRRGSGSARGVRRGARGAQAQQRGRQGVAGPDLQ
 BASE COUNT 699 a 402 c 561 g 559 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,41e-130 Length: 2221
 Score: 2964.50 Matches: 559
 Percent Similarity: 99.82% Conservative: 2
 Best Local Similarity: 99.47% Mismatches: 0
 Query Match: 99.41% Indels: 1
 DB: 9 Gaps: 1
 US-09-821-687-4 (1-561) x AFI55568 (1-2221)
 QY 1 MetAlaThrGlnHisValAsnGlyThrGlnGluProMetAspThrSerAla 20
 DB 205 ATGGCTACAGAACATGTTATATGGAATGCTAAGAGGCCATGATTAATCTTCTGCA 264
 QY 21 ValIleHisSerGlnAsnPheGlnThrLeuAspAlaGlyLeuProGlnLysValAla 40
 DB 265 GTTATTCATTCAGAAATTTTCAGACATTCCTGATGCTGCTTTCACAGAAAGTGTCT 324
 QY 41 GlnLysLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60
 DB 325 GAAATCTGATGAAATTTTACCTTGAGGGCTAGTGGACATGATGATTAATGAAAGA 384
 QY 61 AlaIleGlnAlaLeuLysGlnPheAsnGluAspGlyAlaLeuAlaValLeuGlnInphe 80
 DB 385 GCTATTGAAAGCTTTTAAAGAAATTCATGAAAGCGGTGATTCGACGTTCTTCAACAGTTT 444
 QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuGlyValIleLys 100
 DB 445 AAGACAGGATCTCTCATGATTCAGAACAAAGTCTTTTATGAGTCAATGAG 504
 QY 101 ThrTyrArgGlnArgGlnLysGlnGlyThrLysValAlaAspSerSerLysGlyProAsp 120
 DB 505 ACTTACAGGACAGAGGAAACAGAGGACCAAGATTCATGATTAATGAAAGCAAGAT 564
 QY 121 GlnAlaLysIleLysAlaLeuLeuGlnArgThrGlyTyrThrLeuAspValThrThrGly 140
 DB 565 GAGGCAAAATTAAGGACCTCTGGAAAGAACAGGCTACACCTTGATGTGACCACTGGA 624
 QY 141 GlnArgLysTyrGlyGlyProProProAspSerValTyrSerGlyGlnGlnProSerVal 160
 DB 625 CAGAGGAGATATGAGAGGACCACTCAGATTCCTGTTATTCAGTCAAGACCTCTGTT 684
 QY 161 GlyThrGlnIlePheValGlyLysIleProArgAspLeuPheGlnAspGlnLeuValPro 180
 DB 685 GGCACCTGAGATATTTGTGGAAAGATCCCAAGAGATCTATTGAGAGATGAACCTTGTCCA 744
 QY 181 LeuPheGlnLysAlaGlyProIleThrAspLeuAspLeuMetMetAspProLeuThrGly 200
 DB 745 TTATTGAGAAAGCTGAGACCTATGAGATCTCTGCTATATGATGATCACTCACTGCT 804
 QY 201 LeuAsnArgGlyTyrAlaPheValThrPheGlyThrLysGlnAlaGlnLysValAla 220
 DB 805 CTCATATAGAGGTTATGCGCTTGTCACTTTTGTACAAAAGAGCAGCTCAGAGGCTGTT 864
 QY 221 LysLeuTyrAsnAsnHisGlnLysSerGlyLysHisIleGlyValCysIleSerVal 240
 DB 865 AACTGTATATATCATGAAATGCTTGTGAAACATATGGTGTGCAATCTCAGT 924
 QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnLysLeu 260
 DB 925 GCACCAATATAGGCTTTTGTGGCTCATCTTCTTAAAGATGAACAGAGCAAGTTCTT 984

OY	261	Gluc1pHesSerValThrGluGlyLeuThrAspValIleuGlyTrhisGlnProAsp	280
Db	985	GAGAAATTATGACCAACTAACAGAGGGCTTTACAGAGTCAATTTATACCACAACCGCAT	1044
OY	281	AapLybLYeAbaNarGlyPheCySPheLeuGluTyTrGluaspHisLysThrAlaAla	300
Db	1045	GACCAAAAAAAAAACAGAGGCTTTTGCTTCCTGAATTAAGAATCACAAAACAGCTGCC	1104
OY	301	GlnAlaArgArGArgLeuMetSerGlyLysValLysValITrGlyAsnValGlyThrVal	320
Db	1105	CAGCAGGCGTAGGTAAATGATGATGAAGTCAAAGCTCGGGGAGATGTGAACTGTT	1164
OY	321	GluTrPAIAAsPProIIeGluAsPProAsPProGluIVaIMetAlaLysValLysValLeu	340
Db	1165	GAAATGGCGCTGATCCCTATAAGATCCTGATCCTGAGGTTATGGCAAAGTAAAGCTGC	1224
OY	341	PheValArgAsnLeuAlaAsnThrValIThrGluGluIleLeuGluLysSerPheSerGln	360
Db	1225	TTTGTAACGCACACTTGCCCAATACCTGTAACAGAAAGATTTTGGAAAAAGCATTATGTCAG	1284
OY	361	PheGlyLysLeuGluAspArgValLysLysLeuLysAspTyTrAlaPheIleHisPheAspGlu	380
Db	1285	TTTTGGAAATCTGAAACAGAGTGAAGAAAGTTAAAGATTATGCGTTCATCAATTTGTATGAG	1344
OY	381	ArgAspGlyAlaValLysAlaMetGluGluMetAsnGlyLysAspLeuGluGlyAsn	400
Db	1345	CGAAGATGGTGCTGTCAAGGCTATGAGAAAGATGAATGGCAAGAACTTGGAGGGAGAAAT	1404
OY	401	IleGluIleValPheAlaLysProPAspGlnLysArgLysGluArgLysAlaGlnArg	420
Db	1405	ATTGAATATGTTTTTGCCCAAGCCACACAGATCCAGAAAGAAAGAAAGAAAGCTCAGAGG	1464
OY	421	GlnAlaAlaLysAsnGluMetTyTrAspAspTyTrTyTrTyTrGlyProProHisMetPro	440
Db	1465	CAGCAGCAAAATTCMAATGATGACATTACTACTATTATATGCTCACCTCATATGCCCC	1524
OY	441	ProProThraArgGlyARgGlyARgGlyGlyARgGlyGlyTYTrGlyTYTrProProAspTyTr	460
Db	1525	CCTCCACAAGAGGTGAGGGGGGTGAGGTTAGAGGTGTTATGAGATATCTCCACAAATTAT	1584
OY	461	TyTrGlyTYTrGluAspTyTrTyTrAspTyTrGlyTYTrAspTyTrHisAsnTyTrArgGlyGly	480
Db	1585	TATGATATATGAAGATTATATGATATTAATGTTATGATTACCATTAATATCGTGGTGA	1644
OY	481	TyrGluAspProTyTrTyTrGlyTyTrGluAspPheGlnValGlyAlaArgGlyARgGlyGly	500
Db	1645	TATGATGATCCATTAATGATGATGATGAAATTTTCAAGTTGAGACTTAGAGAAAGGGGTGAT	1704
OY	501	ArgGlyAlaArgGlyAlaAlaLysProSerArgGlyArgGlyAlaAlaLysProProArgGlyArg	520
Db	1705	AGAGGAGCAAGGGGTGCTCTCATTCACAAGGTCGAGGGGCTGCTCTCCCGCGGTAA	1764
OY	521	AlaGlyTYTrSerGlnInArgGlyGlyProGlySerAlaArgGlyValAArgGlyAlaArgGly	540
Db	1765	GCCGGTTATTCACAAGAGAGAGGTCTTGATCAGCAAGAGGGGCTTCAGAGTGCGAGAGGA	1824
OY	541	GlyAlaGlnGlnGlnInArgGlyValArgGly---GlyLysGlyValGlnAlaGlyProAspLeu	559
Db	1825	GGTGGCCCAACACAAAGAGGCGCGGCGCAGGAGAAAGGGGTGAGAGCCGATCTGACTTG	1884
OY	560	LeuGln 561	
Db	1885	TTACAA 1890	
<hr/>			
RESULT 3			
LOCUS	AY034483	1686 bp	mRNA
DEFINITION	Homo sapiens hnRNP Q1 mRNA, complete cds.	linear	FRI 03-OCT-2001
ACCESSION	AY034483		
VERSION	AY034483.1	GI:1580589	
KEYWORD			
SOURCE	Homo sapiens.		

ORGANISM	Hom sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (bases 1 to 1686)
JOURNAL	Moutrelatos, Z., Abel, L., Yong, J., Kataoka, N. and Dreyfuss, G.
MEDLINE	SNM interacts with a novel family of hnRNP and spliceosomal
PUBMED	proteins
REFERENCE	EMBO J. 20 (19), 5443-5452 (2001)
AUTHORS	2 (bases 1 to 1686)
JOURNAL	Moutrelatos, Z.
TITLE	Direct Submission
FEATURES	Submitted (11-MAY-2001) HMMI/ Biochemistry, University of
SOURCE	Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA
	Location/Qualifiers
	1..1686
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	1..1686
	/codon_start=1
	/product="hnRNP Q1"
	/protein_id="AAK59705.1"
	/db_xref="GI:158093530"
	/translation="MATEHVNGNGTEBPMDTSAVHSENPQTLLDGLPQKAEKID
	ELIVAYVAHSDLDIERAIEALKEFNEGALAVIQFSDLSHVONKSAFLCGWKTY
	ROBEQGTIVADSDKSGPDEAKIKALERTYTLTDVGKRGGPPDSVYSGQPSV
	GTELEFKIIPDLFDEDELVLPFKAGITMDLRIMMDPLTGNGYAFVTECKEAOE
	AVLVNHNHRSKGKIGVCI SVANNRLFVSGI IKSKEGQILIEEFSVTEGLTDVILY
	HODDKKKNSPCLEVEDEHKTAAQAPRMSKRYVMGQVETEMDPIEDDPBYM
	AKRYVLEFVRILANTVTEILIEKAPSQGKLERKKLDYAFIHFDEEDGAVKAEENK
	GKQEGENIEIVFAKPEPDQKREKAKORAKOMDDYIYGEPPHNPPTRGGRGGK
	RGGEYPPDYEDYDDYGYDYHNYNRGQYEDPYEDFOVARGRGGRGARGAAP
	SRRGGAAPPRGRAGYSQSGRGSGARVGARGAQQQGRGVGVGEVAGPDLLQ"
BASE COUNT	536 a 282 c 435 g 433 t
ORIGIN	
Alignment Scores:	
Score:	1,12e-130
Percent Similarity:	99.64%
Best local Similarity:	99.47%
Query Match:	99.40%
DB:	9
	Gaps: 0
US-09-821-687-4 (1-561) x AY034483 (1-1686)	
1	MetalaThrgluHsiValaSnGLYAsnGLYThrgluGluGluPROMetAAspThrThrsErala 20
1	ATGCGTACAGAACATGTTAAATGCAAAATGTAAGAGCCCATGATCACTTCTTGCA 60
21	ValIehisSergluAenPhGlnThrleuLeuAapalagIleuProGlnLyValAla 40
61	GTATTCATTGACAAATATTTTCAGACATTCGTTGATGCTGTTATCCACAGAAAGTGCT 120
41	GIuLySerAAspGluIleTyrValAlaGlyLeuValAlaHisSerAAspLeuAAspGluabg 60
121	GAATAACTAGATGAATATTCATGTCAGGCGCTAGTGCACATAGATTAAGTAAAGA 180
61	AlaIleGluAlaLeuLySgluPhAsnGluAAspGlyAlaLeuAlaValleuGlnGlnPhe 80
181	GCTATGAGAGCTTTAAAGATTCATATGAAACGCGTGCATTCGCAAGTTCTTCAACAGTTT 240
81	LySaAspSerAAspLeuSerHisValGlnAenLySerAlaPheLeuCySgLyValMetLyS 100
241	AAAGCACTGATCTCTTCATCATGTCAGAACCAAAAGTGCTTTTATGSGAGTCATGAG 300
101	ThrTyrAAspGlnAAspGluLySglnGlnGlyThrTyrValAlaAspSerSerLySgLyProasp 120
301	ACTTACAGGCGAGAGGAAACAGAGGACCAAGTAGCAATTTCTTAGTAAAGACCAAGAT 360
121	GIuAlaLySgluLySgluAlaLeuGluGluAAspGlnGlyTyrThrLeuAAspValThrThrGly 140

QY 1 MetAlaThrGluHfValAsnGlyAsnGlyThrGluGluProMetAspThrThrSerAla 20
 DB 1 ATGGCTACAGAACTGTTAAATGGAATGCTAGTGAAGGCCATGATCTACTTCTGCA 60
 QY 21 ValIleHisSerGluAsnPhgGlnThrLeuLeuAspAlaGlyLeuProGlnLysValAla 40
 DB 61 GTTATCATTCAGAAATTTTTCAGACATTCCTGATGCTGATTACACAGAAAGTGGCT 120
 QY 41 GluLysLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60
 DB 121 GAATAACTAGTGAATTAATTCAGTTCAGGGCTAGTGCATAGTGAATTAAGTGAAGA 180
 QY 61 AlaIleGluAlaLeuLysGluPhaAsnGluAspGlyAlaLeuAlaValAlleuGlnGlnPhe 80
 DB 181 GCTATTGAAGCTTTAAAGAAATTCATGAGACAGCGTGAATGGCAGTTCTTCAACAGTTT 240
 QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuGlyValMetLys 100
 DB 241 AAAGACAGTGAATCTCTCATGTTCAAGAACAAAGTGCCTTTTATGAGAGCATGAG 300
 QY 101 ThrTyrArgGlnArgGluLysGlnGlyThrLysValAlaAspSerLysGlyProAsp 120
 DB 301 ACTTACAGGACAGAGAGAAACAAAGGACCAAGTACAGATTCATGTAAGACCAAGAT 360
 QY 121 GluAlaLysIleLysAlaLeuLeuGluLysArgThrGlyTyrThrLeuAspValThrThrGly 140
 DB 361 GAGGCAAAATTAAGCACTCTTGGAAGAAACAGGCTTCAACATGATGACCACTGGA 420
 QY 141 GlnArgLysTyrGlyGlyProProAspSerValTyrSerGlyGlnGlnProSerVal 160
 DB 421 CAGAGCAAGTATGAGGACCACTCCAGATTCCTTATTCAGAGTCAGACCTCTGTT 480
 QY 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
 DB 481 GGCCTAGAGATTTGTTGGGAAAGATCCCAAGAGATTCATTTGAGAGTGAACCTGGTCCA 540
 QY 181 LeuPheGluLysAlaGlyProIleTyrAspLeuArgLeuMetAspProLeuThrGly 200
 DB 541 TTATTTAGAAAGCTGACCTCATATGGATCTTCTGCTAAAGATGATCCACTCACTGCT 600
 QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGluAlaAlaGlnGluAlaVal 220
 DB 601 CTCATATAGAGTTATGGCTTGTGCTCATCTTTGTACAAAGAGAGCTCAGAGAGCTGTT 660
 QY 221 LysLeuTyrAsnAsnHisGluIleArgSerGlyLysHisIleGlyValCysIleSerVal 240
 DB 661 AAATCTGATATATATCATGAAATTCGTTCTGAAACATATGCTGATCTCAGTT 720
 QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGluGlnIleLeu 260
 DB 721 GCCAACAATAGGCTTTTGTGGGCTTATTCCTTAAGGCTAAACCAAGAACCAATCTT 780
 QY 261 GluGluPheSerLysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnProAsp 280
 DB 781 GAGAAATTTAGCAAGTAAACAGAGGCTTCAAGACGTCATTTATATCCACCAACCGGAT 840
 QY 281 AspLysLysLysAsnAsnArgGlyPheCysPheLeuGluLysArgHisLysThrAlaAla 300
 DB 841 GACAAAGAAAACAGAGGCTTTTGTCTTCTGTAATATGAAATCAAAAACAGCTGCC 900
 QY 301 GlnAlaArgArgArgLeuMetSerGlyLysValIleLysValIleProLysAsnValGlyThrVal 320
 DB 901 CAGGCAAGGCTAGGTTATAGTGTAAATCAAGGCTCTGGGGAAATGTTGGAACGTT 960
 QY 321 GluTyrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysValLeu 340
 DB 961 GAATGGGCTGATCTATAGAAATCTGATCTGAGGTTATAGCAAAAGTGAAGTCTG 1020
 QY 341 PheValIleArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSerGln 360
 DB 1021 TTTTATGCAAACTTGTCCAAATACGTTAAACAGAAAGATTTTGAAGAAAGCACTTTATGCTAG 1080
 QY 361 PheGlyLysLeuGluArgValLysLeuLysAspTyrAlaPheIleHisPheAspGlu 380

DB 1081 TTTGGAAACTGGAACAGAGTGAAGAAAGTTAAAGATTAATGCGTTACTTATTTGATGAG 1140
 QY 381 ArgAspGlyAlaValLysAlaMetGluGluMetAsnGlyLysAspLeuGluGlyLeuAsn 400
 DB 1141 CGAATGGTGTCTCTCAAGCTATGGAAGAAATGAATGAAGCAACTTGAAGGAGGAAGAAAT 1200
 QY 401 IleGluIleValPheAlaLysProProAspGlnLysArgLysGluLysArgLysValAlaArg 420
 DB 1201 ATGAAATGTTTTCACCAAGCCACAGATCCAGAAAGAAAGAAAGAAAGAAAGCTCAGAG 1260
 QY 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrProProHisMetPro 440
 DB 1261 CAAGACAGCAAAAATCAAAATGATGACGATTTACTATTAATGCTCACCTCATGATGCC 1320
 QY 441 ProProThrArgGlyLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 460
 DB 1321 CTTCCAAAGAGAGTCAAGGCGTGAAGGTAGAGGTGATGATATCCCTCCAGATTTAT 1380
 QY 461 TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspDyrHisAsnTyrArgGlyGly 480
 DB 1381 TATGATATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
 QY 481 TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValAlaArgGlyLysGlyLys 500
 DB 1441 TATGAAGATCCATCATGATGATTAAGAAATTTTCAAGTTGAGCTGAGAGGAGGCTGGT 1500
 QY 501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyLysGlyAlaAlaProProArgGlyLys 520
 DB 1501 AGAGAGCAAGGAGTGTCTGCTCATCCAGAGGTGCGGGGCTGCTCTCCCGCGGTAGA 1560
 QY 521 AlaGlyTyrSerGlnArgGlyGlyProGlyLysArgLysValAlaArgGlyLysArgGly 540
 DB 1561 GCCGTTATTCAGAGAGAGAGGCTCTGATTCAGAAAGGCGTTGAGGTGCCAGAGGA 1620
 QY 541 GlyAlaGlnGlnGlnArgGlyLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 556
 DB 1621 GGTGCCCAACAAAGAGGCGGCGGTACGTGCGGAGGCGGTGC 1668
 RESULT 5
 AF037448 2932 bp mRNA linear PRI 08-APR-1998
 LOCUS Homo sapiens RRM RNA binding protein Gry-rbp (GRY-RBP) mRNA,
 DEFINITION complete cds.
 ACCESSION AF037448
 VERSION AF037448.1 GI:3037012
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 2932)
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.
 Cloning of a RRM RNA binding protein with glycine, arginine and
 tyrosine-rich C-terminus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2932)
 Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.
 JOURNAL Direct Submision
 TITLES Submitted (09-DEC-1997) Department of Biochemistry and Molecular
 Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,
 Beijing 100005, P.R. China
 FEATURES
 source
 1..2932
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene
 1..2932
 /gene="GRY-RBP"
 CDS
 526..2397
 /gene="GRY-RBP"
 /note="RRM RNA binding protein; RNA recognition motif
 binding protein; Glycine, arginine and tyrosine-rich at
 C-terminus"


```
/codon_start=1
/product="Grp-78p"
/protein_id="AA012926.1"
/db_xref="GI:3037013"
/translation="MATEHYNNGTEEPMDTTSVAHSNFTLLDAGLPQVAKELD
EIVAGLVASHSDIDERAILAKFPNEDGALAVIQPKDSLHVNQKSFLOCVMKTY
RQREKGTIVADSSSKGDEAKI KALLERGTLDVTYGRKYGPPDPDSVGGQPSV
GFEIYVWKIPRLFEDELVPLEKAPIMDLRLMDPLTGNRGVAFVETCKEAOE
AVLYNNHEIRSGKIHGVCI SVANNRL FVGSIPKSTKEQILIEEPEKTEGLDVTLY
HOPDKKRSRPFLEVEDHKTAAQARRRLMSKRYVMNVCVENADPIEDPDEVM
AKYKULPYRNLANTTVEELERAFSOFGLERVKIKDYAFIHPDRDAVAMEEMN
GKDLBESNEIIVFAPKPDGRKRERKAQRAANQMDYDYGYGAPMRPPTKGRGAP
RGQYGPDPYEGYEDYDYGYDYNHRYRGYEDPYGYGAPVGAAGRGGRGAGAP
SRGRGAPRGGRGAGYSGRGPGSARGVGAQOQGRGVRGARGRGAGRGVGRK
ADGVNPDKRKQTNQNWGSDPIAQOPLQGGDHSGNRYKSNQEFYDPTFGQMK"
BASE COUNT      868 a      520 c      823 g      721 t
ORIGIN
Alignment Scores:
Pred. No.:      1,82e-128      Length:      2932
Score:          2922.00      Matches:      549
Percent Similarity: 99.10%      Conservative: 2
Best Local Similarity: 98.74%      Mismatches: 5
Query Match:      97.99%      Indels:      0
Gaps:            0
US-09-821-687-4 (1-561) x AF037448 (1-2932)
QY      1 MetAlaThrGluHisValaAngIyaAngIyThrGluGluProMetAspThrThrSerAla 20
Db      526 ATGGCTACAGAAACATGTTATGGAATGTGACGAGAGCCATGATCTACTTCGCA 585
QY      21 ValIleHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeuProGlnIlyValaIa 40
Db      586 GTTATTCATTCAGAAATTTTCAGACATTCCTTGATGCTGTTTACCAAGAAAGTTC 645
QY      41 GluIlyLeuAspGlnIlyIyValaIaGlyLeuValaIaHisSerAspLeuAspGlnIy 60
Db      646 GAAAACTGATGAAATTTTACGTTGCGAGGCTAGTTCATATGATTTAGTGAAGA 705
QY      61 AlaIleGluAlaLeuIySegIupheAsnGluAspGlyAlaLeuAlaValleuGlnIuphe 80
Db      706 GCTATTGAAGCTTTAAAGAAATCAATGAAGACGTCGATTCGACGTTCTTCAACGTTT 765
QY      81 LyAspSerAspLeuSerHisValaGlnAsnIySerAlaPheLeuGlyValaIleIyS 100
Db      766 AAGACAGATCTCTCTCATGTTTCAGAACAAAGTCCCTTTTATGTGAGATCAATGAAG 825
QY      101 ThrIyArgGlnArgGluIySegIyThrIyValaIaAspSerSerIySegIyProAsp 120
Db      826 ACTTTCAGGCAAGGAGAAAAACAAGGACCAAGTACAGATTTCTAGTAAAGACAGAT 885
QY      121 GluAlaIyIleIyAlaLeuLeuGlnIyArgIyThrIyThrLeuAspValIyThrArgIy 140
Db      886 GAGGCAAAATTTAAGCAGCTCTTGAAAGAACAGCTACACCTGATGTGACACATCGGA 945
QY      141 GlnArgIyIyGlyIyProProProAspSerValIySerGlyGlnIyProSerVal 160
Db      946 CAGAGGAGATGAGAGACCACTCCAGATTCGTTTATTCAGATCAGCAGCTTCTGTT 1005
QY      161 GluIyThrGluIlePheValGlyIyIleProArgAspLeuPheGlnAspGluLeuValPro 180
Db      1006 GGCACAGATATTTTGTGGAAAGATCCCAAGATCTATTATGATGAATCACTTGTCCA 1065
QY      181 LeuPheGlnIyAlaGlyProIleTPAspLeuArgLeuMetAspProLeuThrArgIy 200
Db      1066 TTAATTGAAAGAGCTGACCTATATGAGATCTTCTGCTATATGATGATCACTCATGGT 1125
QY      201 LeuAsnArgGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 220
Db      1126 CTCATATAGAGATATCGTTTGTACATTTTGTACAAAAGAACACCTCAGAGGCTGTT 1185
QY      221 IyIleuIyIyAsnAsnHisGlnIleArgSerGlyIyHisIleGlyValaCysIleSerVal 240
```

```
Db      1186 AAACCTATATATATATCATGAAATTCGTTCTGGAAGAACATATGCTGCTCATTCAGTT 1245
QY      241 AlaAsnAsnArgLeuPheValGlySerIleProIySerIyThrIySegIuGlnIleLeu 260
Db      1246 GCCAACAAATAGGCTTTTGTGGGCTCTATTTCTTAAGATGAAGAACCAAGACAGATTCCT 1305
QY      261 GluGluPheSerIyValIyThrGluIyLeuThrAspValIleLeuIyIyIyIyIyIyIy 280
Db      1306 GAAAGATTTAGCAAGATTAACAGAGGCTTACAGACGTCAATTTTATACCAACACCGAT 1365
QY      281 AspIyIyIyIyIyAsnArgGlyIyPheCysPheLeuGluIyIyGluAspHisIyIyThrAlaIa 300
Db      1366 GACAAAGAAAAAACAAGAGCTTTTCTCTTCTGATATAGAGATACCAAAACAGTCCG 1425
QY      301 GlnAlaIyArgArgLeuMetSerGlyIyIyValIyIyIyIyIyIyIyIyIyIyIyIyIy 320
Db      1426 CAGGCAAGGCTAGGTTAATGAGTGTAAAGTCAAGTCTGGGGGAGATGTGGAACTGTT 1485
QY      321 GluTPAlaAspProIleGluAspProAspProGluValMetAlaIyValIyValleu 340
Db      1486 GAATGGCTGATCTCTATAGAGATCTGATCTGAGGTTATGCAAGGTTAAAGTCTG 1545
QY      341 pheValArgAsnLeuAlaAsnThrValIyThrGluGluIleLeuGluIyIySerPheSerGln 360
Db      1546 TTTGTACGCACTTCGCCAATCTGTACAGAGAGATTTTGAAGAGCATTTTACTGAG 1605
QY      361 PheGlyIyIyLeuGlnArgValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 380
Db      1606 TTTGGGAAACCTGAAACGATGGAAGAAATTAAAGATTAAGCTTCATTTTGTAGAG 1665
QY      381 ArgAspGlyAlaValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 400
Db      1666 CAGAGATGCTGCTCAAGGCTATGAGAAAGAAATGAGAAAGACTTGAAGGAGAGAAAT 1725
QY      401 IleGluIleValIyPheAlaIySProProAspGlnIyIyIyIyIyIyIyIyIyIyIyIy 420
Db      1726 ATTTGAATTTGTTTTCAGGACCAACAGTCAAGAAAGAAAGAAAGAAAGCTCAGAG 1785
QY      421 GlnAlaIyIyAsnGlnMetIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 440
Db      1786 CAGGACGCAAAATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845
QY      441 ProProThrArgIyIyArgIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 460
Db      1846 CCTCAACAAAGGCTCAGAGGCTGAGGATGAGGATGAGGATGATGATGATGATGATGAT 1905
QY      461 TyrGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 480
Db      1906 TATGATATGAGATTTATTTATATATATATATATATATATATATATATATATATATAT 1965
QY      481 TyrGluAspProIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 500
Db      1966 TATGAAGATCATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
QY      501 ArgGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 520
Db      2026 AAGAGAGCAAGGGGCTGCTCATTCAGAGGTCGAGGCTGCTCCCGCGGTAGA 2085
QY      521 AlaGlyIyIySerGlnArgGlyIyIyProGlyIySerAlaIyIyIyIyIyIyIyIyIyIy 540
Db      2086 GCCGCTTATTCACAGAGAGAGGCTCTGATGATGACAGAGGCTTCAGAGTTCGAGAGGA 2145
QY      541 GluIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 566
Db      2146 GGTGCTCAACAAAGAGGCGCGGGGTACGTTGTCAGAGGCTGCTG 2193
RESULT 6
AF093821
LOCUS      AF093821      2373 bp      mRNA      linear      ROD 04-OCT-1998
DEFINITION Mus musculus RRM RNA binding protein GRV-RBP mRNA, complete cde.
ACCESSION AF093821
VERSION AF093821.1 GI:3694985
```


Db 1507 AGGGGTGTAGAGAGAGAGGGGTGTCTCCATCCAGAGGTGTGGGTGCTCCCTCC 1566
Qy 518 ArgGlyArgAlaGlyTyrSerGlnArgGlyValProGlySerAlaArgGlyValArgGly 537
Db 1567 CGGTGTAGAGCCGGTTATTCACAGAGAGAGCCCTGGATCAGCAAGAGCGTTCGGGT 1626
Qy 538 AlaArgGlyValAlaGlnGlnArgGlyValArgGlyValGlyValAlaGlyValAlaGly 556
Db 1627 GCGAGAGAGAGTCC---CAACAAAGAGCGCGGGGTACGTGTGCAAGGGGTGCG 1680
RESULT 7
ARI77851 2079 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 2 from patent US 6313266.
DEFINITION ARI77851
ACCESSION ARI77851
VERSION ARI77851.1 GI:11920206
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Bandman, O., Yue, H., Corley, N. C. and Shah, P.
TITLE Human nucleolin-like protein
JOURNAL Patent: US 6313266-A 2 06-NOV-2001;
FEATURES
source 1.2079
location/Qualifiers
BASE COUNT 639 a 377 c 535 g 528 t
ORIGIN
Alignment Scores:
Pred. No.: 9,496-127 Length: 2079
Score: 2882.00 Matches: 547
Percent Similarity: 97.51% Conservative: 1
Best Local Similarity: 97.33% Mismatches: 1
Query Match: 96.65% Indels: 13
DB: 6 Gaps: 1
US-09-821-687-4 (1-561) x ARI77851 (1-2079)
Qy 1 MetalThrGluHisValaGlnGlyAsnGlyThrGluGluProMetAspThrThrSerAla 20
Db 203 ATGGCTACAGAACATGTTATGGAAGTGTACTGAGAGAGCCAGTACTTCTTCTGCA 262
Qy 21 ValIleHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeuProGlnIleValAla 40
Db 263 GTTATCCATTCAAGAAATTTTCAGACATTCCTGATGCTGTTTACACAGAAAGTTGCT 322
Qy 41 GluIleLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60
Db 323 GAAAACTGATATAAATTTCGTGACGGGCTAGTGCACATGTGATTTAGATGAAAGA 382
Qy 61 AlaIleGluAlaLeuLysGluPheAsnGluAspGlyValAlaLeuAlaValLeuGlnGlnPhe 80
Db 383 GCATTTGAAGCTTTAAAGAAATTCATGAGAGACGGTCAATTTGGCAGTTCTTCAACAGTTT 442
Qy 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuGlyValIleLys 100
Db 443 AAAGACAGTATCTCTCTCATGTTTCAGAACAAAGTCCCTTTTATGTGAGCATGAG 502
Qy 101 ThrTyrArgGlnArgGluLysGlnGlyThrIleValAlaAspSerSerLysGlyProAsp 120
Db 503 ACTTACAGGAGAGAGAGAAACAAAGGACCAAGTACAGATCTTCAATAAGACCAAGAT 562
Qy 121 GluAlaLysIleLysAlaLeuGluArgGluArgThrGlyTyrThrLeuAspValThrThrGly 140
Db 563 GAGGCAAAATTTAAGGACCTTGGAAAGAACAGCTACACCTTATGTGACCACTGGA 622
Qy 141 GlnArgLysTyrGlyGlyProProAspSerValTyrSerGlyGlnGlnProSerVal 160
Db 623 CAGAGAGAGTATGAGAGACCACTCCAGATTCGTTTATTCAGGTCCAGACCTTCTGTT 682

Qy 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
Db 683 GGGCTGAGATATTTGTGGAAAGATCCCAAGAGATCTTTGAGGATGACCTTGTCCA 742
Qy 181 LeuPheGluLysAlaGlyProIleTyrAspLeuArgLeuMetAspProLeuThrGly 200
Db 743 TTATTTGAGAAAGCTGACCTATATGAGATCTTCGTTATGATGATGATCACTCATGGGT 802
Qy 201 LeuAsnArgGlyTyrValAlaPheValThrPheCysThrIleGluAlaAlaGlnGluAlaVal 220
Db 803 CTCATATGAGGTTATGCGTTTGTCACTTTTGTATCAAAAGAACAGCTCAGAGGCTGTT 862
Qy 221 LysLeuTyrAsnAsnHisGluIleArgSerGlyLysHisIleGlyValCysIleSerVal 240
Db 863 AATCTGATATATATATATGAAATTCCTTTCGAAACATATGTGTCTGCACTCATGTT 922
Qy 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnGlnIleLeu 260
Db 923 GCCAATATAGCTTTTGTGGGCTCTATTCCTTAAGATTAACCAAGGACAGATTCTT 982
Qy 261 GluGluPheSerLysValThrGlyLeuThrAspValIleLeuTyrHisGlnProAsp 280
Db 983 GAAGATTTAGCAAGCTAACAGAGGCTCTTACAGACGTATTTATACCAACCGCAT 1042
Qy 281 AspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAsnHisLysThrAlaAla 300
Db 1043 GACAAAGAAAGAAACAGAGGCTTTCTTCTTGAATATGAAATATCAAAACAGCTGCC 1102
Qy 301 GlnAlaArgArgArgLeuMetSerGlyLysValIleValTyrGlyAsnAlaGlyThrVal 320
Db 1103 CAGGCAAGGCTATGTTATGATGAGTGAAGTCAAGCTCTGGGGAGATGTGGAACCTGTT 1162
Qy 321 GluThrAlaAspProIleGluAspProAspProGluValMetAlaLysValIleLeu 340
Db 1163 GAATGGCTGATCTCTTACAGAGATCCTGATCCTGAGATTTAGCAAGGTAAGGCTG 1222
Qy 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSerGln 360
Db 1223 TTGTAGGCAACTTCCCAATCTGTACAGAGACATTTTGAAGAAAGCATTTATGTCCAG 1282
Qy 361 PheGlyLysLeuGluArgValIleLysLysLeuLysAspTyrAlaPheIleHisPheAspGlu 380
Db 1283 TTTGGAAATCTGAAAGATGAGAGATTAAGATTAATGCGTCACTTATTTGATGAG 1342
Qy 381 ArgAspGlyValAlaLysAlaMetGluGluMetAsnGlyLysAspLeuGluGluAsn 400
Db 1343 CGAGATGTGTGTCTCAAGGCTATGAAAGAAATGCAATGGCAAGACTTGGAGGAGAAAT 1402
Qy 401 IleGluIleValPheAlaLysProProAspGlnLysArgLysGluArgLysAlaGlnArg 420
Db 1403 ATTGAATGTTTTCCTCAAGCCACCAAGATCAGAAAGAAAGAAAGAAAGCTCAGAG 1462
Qy 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrGlyProProHisMetPro 440
Db 1463 CAGACGCAAAATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1522
Qy 441 ProProThrArgGlyArgGlyArgGlyValArgGlyValTyrGlyProProAspTyr 460
Db 1523 CTTCCAAAGAGGTGAGAGGCGGTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1582
Qy 461 TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGly 480
Db 1583 TATGATATGAAAT 1642
Qy 481 TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValAlaArgGlyArgGlyVal 500
Db 1643 TATGAGATTCAT 1702
Qy 501 ArgGlyAlaArgGlyValAlaAspSerArgGlyArgGlyValAlaAspProAspArgGly 520
Db 1703 AGAGAGCAAGAGGCTGCTCCATCCAGAGGTGAGGAGGCTGCTCCCGCGGTGAG 1762
Qy 521 AlaGlyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValArgGlyAlaArgGly 540

Db 878 GAAGAAATTGACAAAGTAACAGAGGCTTACAGACGTCAATTTATACCAACCAACCGGAT 937
 QY 281 ASPLVLYSLYSANARGGLYPHECYSPHELENGLUYTYRGLUASPHSLYSTRALAA 300
 Db 938 GACAGAGAAAAAACAAGGCTTTTCTTTCTTTGATTAAGATACCAAAACGCTGCC 997
 QY 301 GlnAlaArgArgArgLeuMetSergLyLyValLyValTyrGlyLeuValGlyThrVal 320
 Db 998 CAG----- 1000
 QY 321 GlnTyrAlaAspProIleGluAspProAspProGluValMetAlaLyValLyValLeu 340
 Db 1001 -----GTAAAGTGCTG 1012
 QY 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLySersPheSergLn 360
 Db 1013 TTTGTACGAACTTGCAATCTTCTGTAACAGAGATTTTGAAGAAAGCAATTAGCAG 1072
 QY 361 PheGlyLySleuGluArgValLyLySleuLySAspTyrAlaPheIleHisPheAspGlu 380
 Db 1073 TTTGGGAACTGGAAACGAGAGAGATTAAAGATTATCGTTCAATTTTGATGAG 1132
 QY 381 ArgAspGlyValAlaValLySAlaMetGluGluMetAsnGlyLySAspLeuGluGlyLeuAsn 400
 Db 1133 CGAGATGGTCTGTCAAGGCTATGAGAAATGATGCGCAAGACTTGAGGAGAAAT 1192
 QY 401 IleGluIleValPheAlaLySProProAspGluLySArgLySleuArgLySAlaGlnArg 420
 Db 1193 ATTGAATTTGTTTGGCAAGCCACAGATCAAGAAAGAAAGAAAGAAAGCTCAGAGG 1252
 QY 421 GlnAlaAlaLySAsnGlnMetTyrAspAspTyrTyrTyrGlyProProHisMetPro 440
 Db 1253 CAAGCAGCAAAAATCAAAATGATGAGATTAATCACTATATGCTCACTCAATATGCC 1312
 QY 441 ProProThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 460
 Db 1313 CTTCCAAACAAGAGTGAAGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1372
 QY 461 TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGly 480
 Db 1373 TATGATATGAAATTTATATGATTTATATGATTTATATGATTTATATGATTTAT 1432
 QY 481 TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValGlyAlaArgGlyArgGly 500
 Db 1433 TATGAGATCCATCACTATGTTATGAGATTTTCAAGTTGAGACTGAGAGAGGCTG 1492
 QY 501 ArgGlyValaArgGlyAlaAlaProSeraArgGlyArgGlyAlaAlaProProArgGlyArg 520
 Db 1493 AGAGAGCAAGGAGGCTGCTCCATCCAGAGGTCGTGGGCTGCTCCCTCCCGGTA 1552
 QY 521 AlaGlyTyrSeraGlnArgGlyArgGlySeraAlaArgGlyValaArgGlyValaArgGly 540
 Db 1553 GCGGTTATTCAGAGAGAGAGGCTTCGAGATCAGCAAGAGCGCTGAGGTCGAGAGAG 1612
 QY 541 GlnAlaGlnGlnGlnArgGlyArgGly-----GlyLySleuValGlnAlaGlyProAspLeu 559
 Db 1613 GGGGCCCAACAACAAGAGCGCGCGGCGAGGAGAAAGGGGTCAAGCGGCTCTGACTG 1672
 QY 560 LeuGln 561
 Db 1673 TTACAA 1678
 RESULT 9
 LOCUS AY034482 1767 bp mRNA linear PRI 03-OCT-2001
 DEFINITION Homo sapiens hnRNP Q2 mRNA, complete cds.
 ACCESSION AY034482
 VERSION AY034482.1 GI:15809587
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1767)
 AUTHORS Mourelatos, Z., Abel, L., Yong, J., Katzoka, N. and Dreyfuss, G.
 TITLE SMN interacts with a novel family of hnRNP and spliceosomal proteins
 JOURNAL EMBO J. 20 (19), 5443-5452 (2001)
 MEDLINE 21458434
 PUBMED 11574476
 REFERENCE 2 (bases 1 to 1767)
 AUTHORS Mourelatos, Z.
 TITLE Direct Substitution
 JOURNAL Submitted (11-MAY-2001) HHMI/ Biochemistry, University of Pennsylvania, 330 CRB, 415 Currie Blvd., Philadelphia, PA 19104, USA
 FEATURES
 source location/Qualifiers
 1..1767
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..1767
 /codon_start=1
 /product="hnRNP Q2"
 /protein_id="AA59704.1"
 /db_xref="GI:15809588"
 CDS
 1..1767
 /translation="MATEHVNGNGTEEPMDTTSVAVISENFOTLLDGLPKVAKELD
 EIVYAGLVASDLDERALEALKEFVEDGALVLOQFQSDLSHVONKSAFLCGVAKTY
 ROREKQTKVADS SKGPDEAKI KALERTGYTLDTTGQKXGPPSPVSGQOPEV
 GTEI PVGRI PDLFEDLVPLFEKAGPIWDLRIMMDPLTGIRGYAFTPTCKEAOE
 AVKLYNNHISGKHIGICISVANNRLFVSGIPSKYKSGQILBEFSKTBGLTDTLY
 HOPDKKRNRSFCLEVDHKTAAQVKLFVNNANTVTEIILEAFSQFLERVK
 LKDVAFHIFDRDGAVKAMEBNNGDLSEINLEIFAPPPQKREKRAQOAAKNOM
 YDDVYVYGPPIHPPTRRGRGRGGRGYVPPDYGYEDYDYDYGYDYNHNGYEDPY
 YGEDPVGARGRGGRGARGAAPSRGGAAPPRGAGVSGRGSAGVAGVARGAGAO
 OORGVGVAGRGGRGVNGVGGKRGKADGVNQPSKRGRTNNQWGSPIAQOPLQGGDHS
 GNYGKSENQERTPDTPGQWKK"
 BASE COUNT 561 a 310 c 453 g 443 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,456-118 Length: 1767
 Score: 2705.50 Matches: 514
 Percent Similarity: 92.81% Conservative: 2
 Best Local Similarity: 92.45% Mismatch: 5
 Query Match: 90.73% Indels: 35
 DB: 9 Gaps: 1
 US-09-821-687-4 (1-561) x AY034482 (1-1767)
 QY 1 MetAlaThrGluHisValAsnGlyAsnGlyThrGluGluProMetAspThrSeraLa 20
 Db 1 ATGGCTACAGAAACATGTTATGGAATGTTACTGAAGGCCATGGAATCTCTTCA 60
 QY 21 ValIleHisSerGluAsnPheGlnThrLeuAspAlaGlyLeuProGlnLySAla 40
 Db 61 GTTATCCATTCAGAAATTTTTCAGACATTCCTGATGCTGCTTATACCAAGAAAGT 120
 QY 41 GlnLySLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60
 Db 121 GAAAACTTGATGAATTTTACGTGAGGCTGAGTTCACATGATGATTTAATGAAGA 180
 QY 61 AlaIleGluAlaLeuLySLeuPheAsnGluAspGlyAlaLeuAlaValLeuGlnGlnPhe 80
 Db 181 GCTATTTGAAGCTTTTAAAGATTCATGAGAGCGGTGATTCGAGTTCTTCAACGTTT 240
 QY 81 LySAspSerAspLeuSerHisValGlnAsnLySeraAlaPheLeuCySGLyValMetLyS 100
 Db 241 AAAGACAGTATCTCTCATGTTCAAGCAAAAGTGCCTTTTATGTGAGTCAATGAAG 300
 QY 101 ThrTyrArgGlnArgGluLySglnGlyThrLySValAlaAspSerSeraGlyProAsp 120
 Db 301 ACTTACGCGACAGAGAAAAACAAGGACCAAGTACAGATTTCAAGTGAAGACCGAT 360
 QY 121 GlnAlaLySLeuAlaLeuLeuGluArgThrGlyTyrThrLeuAspValThrThrGly 140
 Db 361 GAGGCAAAATTAAGGCACTTTGGAAAGAACAGGCTACACACTTGATGTACCACTGA 420

QY 141 GlnArgLysTyrGlyProProAspSerValTyrSerGlyGlnGlnProSerVal 160
 DB 421 CAGAGAGAGTATGAGAGCACCCTCCAGATTCGGTTTTCAGGTCCAGAGCTTCCTT 480
 QY 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
 DB 481 GGCCTGATGATATTTGGGAGAAAGATCCAGAGATCTATTTGGAGATGAGACTGGTCCA 540
 QY 181 LeuPheGluValAspIleGlyProIleTyrAspLeuArgLeuMetLeuAspProLeuThrGly 200
 DB 541 TTATTTGAGAAAGCTGAGCTCATATATGATATCTTGGTAAAGATGAGATCCACTAGGT 600
 QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGluAlaAlaGlnGluAlaVal 220
 DB 601 CTCAATAGAGTTATGCTTGTCTTCATTTTGTACAAAGAGAGAGCTCCAGAGGCTGTT 660
 QY 221 LysLeuTyrAsnAsnHisGluIleArgSerGlyLysHisIleGlyValCysIleSerVal 240
 DB 661 AACTGTAATATATCATGAAATTCGTTCTGAAAAACATATGGGTCTGCATCTCAGTT 720
 QY 241 AlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGluGlnIleLeu 260
 DB 721 GCCAACAATAGCTTTTGTGGCTCTATATCTTAAGATTAACCAAGAGAGAGATTCCTT 780
 QY 261 GluGluPheSerLysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnProAsp 280
 DB 781 GAAGATTTTACCAAGTACAGAGGCTCTTACAGCGCATTTTATACCAACAGATTCCTT 840
 QY 281 AspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAspHisLysThrAlaAla 300
 DB 841 GACAGAGAAAAACAGAGAGCTTTGCTTCTGAAATGAAATCAACAAACAGCTGCC 900
 QY 301 GlnAlaArgArgLeuMetSerGlyLysValLysValIleProGlyAsnValGlyThrVal 320
 DB 901 CAG----- 903
 QY 321 GluTyrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysValLeu 340
 DB 904 -----GTAAAAGTCTG 915
 QY 341 PheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSerGln 360
 DB 916 TTGTAGCGCAACTTGCATATCTGTAAACAGAGATTTTGAAGAGAGCTTTTATGCTAG 975
 QY 361 PheGlyLysLeuGluArgValLysLeuLysLysAspTyrAlaPheIleHisPheAspGlu 380
 DB 976 TTTCGGAACCTGGAACGAGTAAAGATTAAGATTAAGCTTCATTCATTTGATGAG 1035
 QY 381 ArgAspGlyAlaValLysAlaMetGluGluMetCysAsnGlyLysAspLeuGluGluAsn 400
 DB 1036 CGAGATGCTGCTCAAGGCTATGGAAGATTAAGTCCAAAGAGCTTCGAGGAGAGAAAT 1095
 QY 401 IleGluIleValPheAlaLysProProAspGlnLysArgLysGluArgLysAlaGlnArg 420
 DB 1096 ATTCGAATTTGTTTTTGCAGAGCCACGATCAGAAAGAGAAAGAGAAAGAGCTCAGAG 1155
 QY 421 GlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrTyrGlyProProHisIleMetPro 440
 DB 1156 CAAAGAGAGAAATCAAGTATGAGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 1215
 QY 441 ProProThrArgGlyArgGlyArgGlyValLysGlyGlyTyrGlyTyrProProAspTyr 460
 DB 1216 CCTCAACAAGAGGTCCAGGCGGTGAGGTAGGTGTTATGATATCCCTCCAGATTTAT 1275
 QY 461 TyrGlyTyrGluAspTyrTyrAspTyrTyrGlyTyrAspTyrHisAsnTyrArgGlyVal 480
 DB 1276 TATGATATGAGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 1335
 QY 481 TyrGluAspProTyrTyrGlyTyrGluAspPheGlnValGlyAlaArgGlyArgGlyVal 500
 DB 1336 TATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 1395
 QY 501 ArgGlyAlaArgGlyAlaAlaProSerArgGlyLysGlyAlaAlaProProArgGlyArg 520

DB 1396 AGAGAGCAAGGGGTGCTCTCCATCCAGAGTCTGGGGCTCTCTCCCGCGTAGA 1455
 QY 521 AlaeGlyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValaArgGlyAlaArgGly 540
 DB 1456 GCCGTTATTCACAGAGAGAGGTCTCGATCAGCAAGAGCGCTTCAGGTGCGAGAGGA 1515
 QY 541 GlyAlaGlnGlnGlnArgGlyArgGlyLysGlyValGluAlaGly 556
 DB 1516 GGTGCCCAACAAAGAGCGCGGAGTACGTGTGCGAGAGGGTGGC 1563
 RESULT 10
 HSJ75C13
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP4-775C13 on chromosome 20p12.1-13. Contains an hnRNP R (heterogenous nuclear ribonucleoprotein R, RRM rna binding protein GR1-RBP) pseudogene, ESTs, STSs and GSS, complete sequence.
 ACCESSION
 AL109618
 VERSION
 AL109618.1 GI:5668665
 KEYWORDS
 HTG; GR1-RBP; hnRNP; RRM.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Direct SubMISSION
 Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requesters: clonerequest@sanger.ac.uk
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr20
 RP4-775C13 is from the library RPCR-4 constructed by the group of Plier de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP4-775C13.
 FEATURES
 source
 1..103819
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="p12.1-13"
 /clone="RP4-775C13"
 /clone_1fb="RPCR-4"
 repeat_region
 1..236
 /note="AluY repeat: matches 1..236 of consensus"
 repeat_region
 293..562
 /note="AluMc repeat: matches 144..1624 of consensus"
 repeat_region
 819..1471
 /note="AluMc repeat: matches 969..1634 of consensus"
 repeat_region
 2177..2588

```

repeat_region /note="LIM3c repeat: matches 14..431 of consensus"
2593..2652
/note="MULTIF repeat: matches 474..533 of consensus"
2701..3176
/note="238 copies 2 mer tt 54% conserved"
2814..2978
/note="15 copies 11 mer 66% conserved"
2996..3196
/note="67 copies 3 mer tcc 56% conserved"
3031..3184
/note="14 copies 11 mer 63% conserved"
3296..4082
/note="LIM4 repeat: matches 3350..4128 of consensus"
4197..4315
/note="LIM4 repeat: matches 3021..3136 of consensus"
4316..4738
/note="MSTA repeat: matches 1..426 of consensus"
4739..5118
/note="LIM4 repeat: matches 2626..3021 of consensus"
5146..5288
/note="match: GSS: Em:B90786"
5345..6619
/note="LIMPA16 repeat: matches 3863..5135 of consensus"
6882..7139
/note="129 copies 2 mer gg 58% conserved"
7351..8391
/gene="du775C13.1"
7351..8391
/gene="du775C13.1"
/note="du775C13.1"
riboprotein R, RRM rna binding protein GRI-RBP)
pseudogene)
match: cDNAs: Em:AF000364 Em:AF037448 Em:AF093821
match: ESTs: Em:AA411147 Em:W38408 Em:AF65128 Em:AV14716
Em:AA11148 Em:AA307374 Em:T89820 Em:AA677426 Em:T89909
Em:AI688219 Em:AI472757 Em:AI548559 Em:AI34844
Em:AA912195 Em:AA094744 Em:AI084864 Em:H34275 Em:AA403140
Em:AA189939 Em:ID6738 Em:AA146046 Em:T39242 Em:AA407580
Em:AA561471 Em:AA647246 Em:AV115211 Em:AA356549
Em:AI581194 Em:H32663 Em:AI566089 Em:AA207215 Em:AA216046
Em:T35450 Em:AA370336 Em:AA655413 Em:W39501 Em:W61969
Em:AA164877 Em:AA403169 Em:AA465343 Em:NE6522 Em:R29647
Em:AA298077 Em:AI625087 Em:AA164209 Em:AA164208
Em:AA563205 Em:AA066208 Em:AA045158 Em:AA524815
Em:AA544174 Em:AA188195 Em:AA64286 Em:AA164876
Em:AI267372 Em:AA130302 Em:AA64286 Em:AA662966
Em:AA130625 Em:AA322633 Em:AA242157 Em:AA130793
Em:AA654252 Em:AA306863 Em:AI186327 Em:AA153325
Em:AA418952 Em:T89722 Em:AA629942 Em:AA95153 Em:AA594467
Em:AV141413 Em:AI469082 Em:AI263768 Em:AA124310
Em:AI072424 Em:AA131602
match: proteins: Tr:O23093 Tr:O24188 Tr:O64985 Tr:O01671
Sw:PP22626 Sw:P38159 Tr:O88569 Tr:O05056 Tr:O88921
Tr:O22388 Tr:O76494 Tr:O50602 Tr:O17309 Tr:O61474
Tr:O22390 Wp:CE02193 Tr:O97560 Sw:D48810 Tr:O43347
Tr:Q24668 Wp:CE03230 Tr:O76181 Tr:D41518 Tr:O43390
Tr:O17310 Sw:Q99020 Sw:PS1968 Tr:Q91579 Wp:CE17059
Sw:PA9314 Sw:O13247 Tr:O04240 Tr:O15717 Tr:O61954
Sw:O08473 Sw:P39697 Tr:O23288 Tr:Q43349 Tr:O81997
Sw:PS1991 Wp:CE00983 Tr:O08937 Tr:P90699 Tr:O15550
Sw:Q01560 Tr:Q24409 Tr:P70372 Tr:O61374 Tr:Q39953
Tr:P70807 Wp:CE02197 Tr:O14979"
/pseudo
/codon_start=1
/evidence=not_experimental
7549..8004
/gene="du775C13.1"
/note="match: GSS: Em:AC434110"
7614..7732
/gene="du775C13.1"
/note="match: GSS: Em:AO536515"
Complement (6065..8562)
/note="match: STRs: Em:T897722"

```

misc_feature	complement(8476..9033)	
	/note="match: GSS: Em:AQ008463"	
repeat_region	8721..8816	
	/note="32 copies 3 mer att 65% conserved"	
repeat_region	9341..10347	
	/note="L1P1A6 repeat: matches 5128..6150 of consensus"	
repeat_region	10381..10517	
	/note="L1M4 repeat: matches 2211..2348 of consensus"	
repeat_region	11105..11201	
	/note="HAI1 repeat: matches 675..775 of consensus"	
repeat_region	11261..11455	
	/note="I2 repeat: matches 645..869 of consensus"	
repeat_region	11456..11765	
	/note="A1uX repeat: matches 1..307 of consensus"	
repeat_region	11766..11863	
	/note="I2 repeat: matches 555..645 of consensus"	
repeat_region	12356..12519	
	/note="I2 repeat: matches 298..464 of consensus"	
repeat_region	12520..12615	
	/note="L1H5 repeat: matches 6051..6146 of consensus"	
repeat_region	13257..13336	
	/note="70 copies 2 mer ga 69% conserved"	
repeat_region	13442..13893	
	/note="MERSB repeat: matches 1..498 of consensus"	
repeat_region	13894..13931	
	/note="19 copies 2 mer ac 89% conserved"	
repeat_region	14475..14910	
	/note="L1M4 repeat: matches 5707..6148 of consensus"	
repeat_region	14960..15480	
	/note="L1M4 repeat: matches 4030..4591 of consensus"	
repeat_region	15498..16282	
	/note="L1P4 repeat: matches 5361..6146 of consensus"	
repeat_region	16286..16427	
	/note="L1P4 repeat: matches 5213..5358 of consensus"	
misc_feature	16428..16958	
	/note="match: GSS: Em:AQ427784"	
repeat_region	17470..18240	
	/note="L1P1 repeat: matches 5277..6151 of consensus"	
repeat_region	18383..18706	
	/note="A1uY repeat: matches 1..309 of consensus"	
repeat_region	18793..18837	
	/note="15 copies 3 mer att 75% conserved"	
repeat_region	19213..19509	
	/note="A1uB repeat: matches 3..298 of consensus"	
misc_feature	complement(join(19771..19905,51843..51912))	
	/note="match: GSS: Em:AQ377847"	
repeat_region	22426..22613	
	/note="MERS3 repeat: matches 4..188 of consensus"	
repeat_region	23385..23551	
	/note="MERSB repeat: matches 1..178 of consensus"	
repeat_region	23821..23871	
	/note="I2 repeat: matches 2445..2496 of consensus"	
repeat_region	23899..24082	
	/note="MERSA repeat: matches 1..181 of consensus"	
repeat_region	24363..24676	
	/note="A1uX repeat: matches 3..312 of consensus"	
repeat_region	25302..25481	
	/note="M1T1G repeat: matches 9..187 of consensus"	
repeat_region	25486..25557	
	/note="26 copies 2 mer ct 86% conserved"	
repeat_region	25564..25591	
	/note="M1T1G repeat: matches 243..270 of consensus"	
repeat_region	26477..26762	
	/note="143 copies 2 mer tc 56% conserved"	
repeat_region	26510..26718	
	/note="19 copies 11 mer 66% conserved"	
repeat_region	26836..27041	
Alignment Scores:		
Pred. No.:	2,266-115	Length: 103819
Score:	2676.00	Matches: 537
Percent Similarity:	95.93%	Conservative: 5
Best Local Similarity:	95.04%	Mismatch: 18

Query Match: 89.74% Indels: 9
 DB: 9 Gaps: 2
 US-09-821-687-4 (1-561) x HSJ775C13 (1-103819)

QY 1 MetAlaThrGluHisValAsnGlyAsnGlyThrGluGluProMetAspThrThrSerAla 20
 DB 7351 ATGGCTACAGAACATGTTAATGAAAATGATCTGAAGAGCCATGATCTACTTCTGCA 7410

QY 21 ValIleHisSerGluAsnPhenGlnThrLeuLeuAspAlaGlyLeuProGlnLysValAla 40
 DB 7411 GTATTCATTCAGAAAATTTTTCAGACATTCCTGATGCTGTTTATCACAGAAAGTTGCT 7470

QY 41 GluLysLeuAspGluIleTyrValAlaGlyLeuValAlaHisSerAspLeuAspGluArg 60
 DB 7471 GAAAACTGATGAAATTTTACGTTGACGGCTAGTGCACATGATTCATTTAGAAAGA 7530

QY 61 AlaIleGluAlaLeuLysGluPheAsnGluAspGlyAlaLeuAlaValLeuGlnGlnPhe 80
 DB 7531 GCTATTCAGCTTTTAAAGATTCATATGACACGGTGCATTCGGCATTTCTTCAACAGTT- 7589

QY 81 LysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeuCysGlyValMetLys 100
 DB 7590 AAAGACAGTATCTCTCTCGTTTCAGAACAAAGTGCCTTTTATGAGATCATGAAA 7649

QY 101 ThrTyrArgGlnArgGluLysGlnGlyThrLysValAlaAspSerLysGlyProAsp 120
 DB 7650 CTTATATGGCAGAGAAAACAAAGGACCAAGTACAGATTTCTAGTAAAGACAGAT 7709

QY 121 GluAlaLysIleLysAlaLeuLeuGlnArgThrGlyTyrThrLeuAspValThrThrGly 140
 DB 7710 GAGGCAAAAATTAAGGCACTTTCGAAAGACAGGCTACACATTTAGTACCACTGCA 7769

QY 141 GlnArgLysTyrGlyGlyProProAspSerValTyrSerGlyGlnProSerVal 160
 DB 7770 CAGAGAAAATATGAGAGCACCTCCAGATTCGGTTTATTCAGGTCAGACACTTCGT 7829

QY 161 GlyThrGluIlePheValGlyLysIleProArgAspLeuPheGluAspGluLeuValPro 180
 DB 7830 GGCACTGATATTTGTTGGGAAAGATCCCAAGATCTATTGAGATGAACTGTTCCA 7889

QY 181 LeuPheGluLysAlaGlyProIleTyrAspLeuArgLeuMetCysAspProLeuThrGly 200
 DB 7890 TTAATTTGAAAGCTGCTCTCATATGGAATCTTCCTCATATGATGATCCACTGCT 7949

QY 201 LeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGlnAlaAlaGlnGluAlaVal 220
 DB 7950 CTCATATAGGCTTATGTTGTTGTCACCTTTTGTAACAAA---GCAGTCGGGAGGCTGTT 8006

QY 221 LysLeuLysTyrAsnAsnHisGlnIleArgSerGlyLysHisIleGlyValCysIleSerVal 240
 DB 8007 AAAACTGATATATATCATATGAAATTCGTTGGAACCAACATGATGTCGATCTCAGT 8066

QY 240 LAlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThrLysGlnGlnIle 260
 DB 8067 TGGCAACATAGGCTTTTGTGGGCTCATATTCCTAAGAGTTAAACCAAGAACAGATTC 8126

QY 260 eugluGluPheSerLysValThrGluGlyLeuThrAspValIleLeuTyrHisGlnPro 279
 DB 8127 TTGAAGAAATTTAGCAAAAGTAAACAGAGGCTCTTACAGACGTCATTTTACACCAACCG 8186

QY 280 AspAspLysLysLysAsnArgGlyPheCysPheLeuGlnTyrGlnAsnHisLysThrAla 299
 DB 8187 GATGCAAGAAAACCA---GGCTTTGCTTTCTTAATATGAAGATCACAAAACAGCT 8244

QY 300 AlaGlnAlaArgArgArgLeuMetSerGlyLysValLysValTyrGlyAsnValGlyThr 319
 DB 8245 GCCCAGGCAAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8304

QY 320 ValGluTyrAlaAspProIleGluAspProAspProGluValMetAlaLysValLysVal 339
 DB 8305 GTTGAATGGGCTGATCTATGAGATGATCTGATCTGATCTGATCTGATCTGATCTGAT 8364

QY 340 LeuPheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGluLysSerPheSer 359
 DB 8365 CTGTTGTAGCAACCTTCGCAATCTGTACGGAAGATTTTACAAAAGGCAATTAAT 8424

QY 360 GlnPheGlyLysLeuGluLysValLysLysLeuLysAspTyrAlaPheIleHisPheAsp 379
 DB 8425 CAGTTGGGAAATGGAATGAGTGAAGAGTTAAAGATTAATCATTCATTCATTTTGGT 8484

QY 380 GluArgAspGlyAlaValLysAlaMetGluGluMetCysAsnGlyLysAspLeuGluGlyGln 399
 DB 8485 GAGTGAATGCTGCTGCTCAAGGCTATGGAAGAAATGATGATGATGATGATGATGAT 8544

QY 400 AsnIleGluIleValPheAlaLysProProAspGlnLysArgLysGluArgLysAlaGln 419
 DB 8545 AATATTGAATTTTGTTCAGACCAACAAATCAAAAAGAAAGAAA---AAGCTCAG 8602

QY 420 ArgGlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrGlyProProHisMet 439
 DB 8603 AGGCAAGACGCAAAAATCAATGATGAGANTTACTATATATGTCACCTTATATG 8662

QY 440 ProProProThrArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 459
 DB 8663 CCCCCCTCAACAAAGATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 8722

QY 460 TyrTyrGlyTyrGluAspTyrTyrAspTyrTyrAspTyrHisAsnTyrArgGly 479
 DB 8723 TATATGATATGATGATGATATATATATATATATATATATATATATATATATATAT 8782

QY 480 GlyTyrGluAspProTyrTyrGlyTyrGluAspPheGlnValGlyAlaArgGlyArgGly 499
 DB 8783 GATATGAATCATCATCATGTTATGAAAGATTTCAAGTTGAGCTGAGAGA-AGGGGT 8841

QY 500 GlyArgGlyAlaArgGlyAlaAlaProSerArgGlyArgGlyAlaAlaProProArgGly 519
 DB 8842 GGTAGAGAACAAAGGGGTGCTGCTCATCCAGAGGTCGTGGGGCTCTCTCCCGTGT 8901

QY 520 ArgAlaGlyTyrSerGlnArgGlyGlyProGlySerAlaArgGlyValArgGlyAlaArg 539
 DB 8902 AGAGCCGTTATTCAG 8961

QY 540 GlyGlyAlaGlnGlnGlnArgGlyArgGly---GlyLysGlyValGluAlaGlyProAsp 558
 DB 8962 GAGAGTCCCAACAAACAAAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9021

QY 559 LeuLeuGln 561
 DB 9022 CTGTTACAA 9030

RESULT 11
 AL390737 212659 bp DNA 1linear HMG 21-MAR-2001
 LOCUS Homo sapiens chromosome 13 clone RP11-190P17, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 16 unordered pieces.
 ACCESSION AL390737 GI:13398796
 VERSION AL390737.9
 KEYWORDS HMG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 212659)
 AUTHORS Burton, J.
 TITLE Direct Submision
 JOURNAL Submitted (17-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Mar 21, 2001 this sequence version replaced gi:13396612.
 COMMENT
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information


```

Center project name: BA190P17
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 203924 bases at least Q40
Consensus quality: 207400 bases at least Q30
Consensus quality: 203306 bases at least Q20
Insert size: 21159; sum-of-contigs
Quality coverage: 7.54x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 48415: contig of 48415 bp in length
* 48416 48515: gap of 100 bp
* 48516 82643: contig of 34128 bp in length
* 82644 82743: gap of 100 bp
* 82744 89101: contig of 6358 bp in length
* 89102 89201: gap of 100 bp
* 89202 127367: contig of 38166 bp in length
* 127368 127467: gap of 100 bp
* 127468 169067: contig of 41600 bp in length
* 169068 169167: gap of 100 bp
* 169168 171259: contig of 2092 bp in length
* 171260 171359: gap of 100 bp
* 171360 173612: contig of 2253 bp in length
* 173613 173712: gap of 100 bp
* 173713 187445: contig of 13733 bp in length
* 187446 187545: gap of 100 bp
* 187546 191084: contig of 3539 bp in length
* 191085 191184: gap of 100 bp
* 191185 194709: contig of 3525 bp in length
* 194710 194809: gap of 100 bp
* 194810 197824: contig of 3015 bp in length
* 197825 197924: gap of 100 bp
* 197925 200623: contig of 2699 bp in length
* 200624 200723: gap of 100 bp
* 200724 203433: contig of 2710 bp in length
* 203434 203533: gap of 100 bp
* 203534 206523: contig of 2990 bp in length
* 206524 206623: gap of 100 bp
* 206624 208875: contig of 2252 bp in length
* 208876 208975: gap of 100 bp
* 208976 212659: contig of 3684 bp in length.
*
* Location/Qualifiers
* 1..212659
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="13"
* /clone_lib="RP11-190P17"
* /clone_1fb="RP11-1.1"
* 1..48415
* /note="assembly_fragment:03097
* fragment_chain:1"
* misc_feature
* 48516..82643
* /note="assembly_fragment:01394
* fragment_chain:1"
* misc_feature
* 82744..89101
* /note="assembly_fragment:01816
* fragment_chain:1"
* misc_feature
* 89202..127367
* /note="assembly_fragment:03032
* fragment_chain:2"
* misc_feature
* 127468..169067
* /note="assembly_fragment:00218
* fragment_chain:2"
* misc_feature
* 169168..171259

```

```

misc_feature /note="assembly_fragment:00015"
171360..173612
/note="assembly_fragment:00477"
173713..187445
misc_feature /note="assembly_fragment:00780"
187546..191084
misc_feature /note="assembly_fragment:01534"
191185..194709
misc_feature /note="assembly_fragment:01673"
194810..197824
misc_feature /note="assembly_fragment:02018"
197925..200623
misc_feature /note="assembly_fragment:02519"
200724..203433
misc_feature /note="assembly_fragment:02839"
203534..206523
misc_feature /note="assembly_fragment:02989"
206624..208875
misc_feature /note="assembly_fragment:03934"
208976..212659
/note="assembly_fragment:04063"
misc_feature /note="assembly_fragment:04063"
BASE COUNT 61246 a 41295 c 42666 g 65947 t 1505 others
ORIGIN
Alignment Scores:
Pred. No.: 4,71e-115 Length: 212659
Score: 2676.00 Matches: 537
Percent Similarity: 95.93% Conservat: 5
Best Local Similarity: 95.04% Mismatches: 18
Query Match: 89.74% Indels: 9
DB: 2 Gaps: 2
US-09-821-687-4 (1-561) x AL390737 (1-212659)
QY 1 MetAlaThrgluHieValaenglyAnglyThrGluGluProMetAspThrSerAla 20
DB 83107 AAGGCTACAGAACATCTTATGAAATGATCTGAAGGCCATGATCTCTTCTGCA 83166
QY 21 ValIleHisSerGluAsnDheGlnThrLeuAspAlaGlyLeuProGlnIysValAla 40
DB 83167 GTATTCATTCAGAAATTTTCAGACATTCGTGATGCTGTTATCACAGAAAGTGT 83226
QY 41 GluIysLeuAspGluIleIysValAlaAglyLeuValAlaHisSerAspLeuAspGluAla 60
DB 83227 GAAAACTGAGTGAATTTACCTGCGAGGCTGATGCAATGATGCAATGATGAAAGA 83286
QY 61 AlaIleGluAlaLeuIysGluIleAspGluAspGluAlaLeuAlaValIleuGlnGlnIle 80
DB 83287 GCTATTGAGCTTTAAAGAAATTCATGAAAGACGGTGATGCGAGTTCTTCAACAGTT- 83345
QY 81 IysAspSerAspLeuSerHisValGlnAsnIysSerAlaPheLeuCySGlyValIleClys 100
DB 83346 AAAGACAGAGATCTCTCTGCTTCAAGAACAAAGCGCTTTTATGTCGATCTATGAA 83405
QY 101 ThrTyrArgGlnArgGluIysGlnGlyThrIysValAlaAspSerSerIysGlyProAsp 120
DB 83406 CCTTATGAGCAGAGAAACAAAGGACCAAGATGACGATTTAGTAAAGACCAAGAT 83465
QY 121 GluAlaIysIleLeuAlaLeuIleuGlnArgThrGlyTyrThrLeuAspValAlaThrThrGly 140
DB 83466 GAGGCAAAATTTAAGGACCTCTTGAAAGACGAGTCAACCTGATGTGACCACTGGA 83525
QY 141 GlnArgIysTyrGlyGlyProProProAspSerValIysSerGlyGlnGlnProSerVal 160
DB 83526 CAGAGGAATATGAGACCACTCAAGATTCGTTTATTCAGGTCAGAGGCTCTCTGTT 83585
QY 161 GlyThrGluIlePheValGlyIysIleProAspLeuPheGluAspGluIleValPro 180
DB 83586 GGCACCTGAGATATTTGTGGAAAGATCCCAAGAGATCTATTTGAGAGATGACCTTGTCCA 83645
QY 181 LeuPheGluIysAlaGlyProIleThrAspLeuAspLeuMetMetAspProIleuThrGly 200
DB 83646 TTAATTGAGAAAGCTGCTCTTATGAGATCTTCCTTAATGATGATCCACTGCTGCT 83705

```


Db 183 GCACCTTTGAAAGAGAGGCTACACACTGATGTGACTACAGGTGACAGAGAAATGGA 242
 Qy 146 G1ProProProAspSerValTyrSerG1yGlnGlnProSerValG1yThnGlnIlePhe 165
 Db 243 GACACCACTCCAGATCCGTTATTCAGGTGACAGGCTTCTGTGGCAGCATATTTT 302
 Qy 166 ValG1yLysIleProArgAspLeuPheG1uAspG1uLeuVal1ProLeuPheG1uLysAla 185
 Db 303 GTGGGGAAGATCCCAAGATCTCTTTGAGATGAGCTTGTTCATTTTGGAAAGCT 362
 Qy 186 G1yProIleTyrAspLeuArgLeuMetMetAspProLeuThnG1yLeuAsnArgG1yTyr 205
 Db 363 GACCTTATAGGATCTTCTGTTATATGATGATCCGCTCACTGCTCAACAGAGGTTAT 422
 Qy 206 AlaPheValThrPheCysThnIyG1uAlaAGlnG1uAlaVal1yLysLeuTyrAsnAsn 225
 Db 423 GCGTTGTCTACTTTTGTACAAAAGAGCAGCACAAGAGGCTGTAAACTGTATATATAT 482
 Qy 226 HisG1uIleArgSerG1yLysHisIleG1yValCysIleSerValAlaAsnAsnArgLeu 245
 Db 483 CATGAATTCGTTCCGGGAAGCAGCATTTGCTGTCTGCACTTCAGTTGCCAACAATAGCTT 542
 Qy 246 PheValG1ySerIleProLysSerLysThnIyG1uGlnIleLeuGlnIleupheSerLys 265
 Db 543 TTTGTGGCTCGATTCCTAAGAGTAAACCAAGAGCAGATTCCTTGAGATTTAGCAA 602
 Qy 266 ValThnGlnIyLeuThrAspValIleLeuTyrHisGlnProAspAspLysLysAsn 285
 Db 603 GTACAGAGGGGTCTCACAGATGTCATTTATTCACCAACCTGTATGCAAGAAAAAAC 662
 Qy 286 ArgG1yPheCysPheLeuG1uTyrG1uAspHisLysThnAlaIleGlnIleArgArgArg 305
 Db 663 AGAGGCTTTGCTTCTTGATATGAGATCACAAACAGCTGCCAGGCAAGCTGAG 722
 Qy 306 LeuMetSerG1yLysVal1yValTyrG1yAsnValG1yThnValG1uTyrPalAspPro 335
 Db 723 CTATAGTGTGTAAGTCAAGTCTGGGAAATGTTGAACTGTGAGTGGGCTGATCCT 782
 Qy 326 IleGlnAspProAspProG1uValMetAlaLysVal1yLysLeuPheValAsnLeu 345
 Db 783 ATTTGAAGATCTTATCTGTAAGTATGAGCAAGGTAAAGTGTCTGTATGACCACTT 842
 Qy 346 AlaAsnThnValThnGlnIleLeuG1uLysSerPheSerGlnPheG1yLysLeuGln 365
 Db 843 GCCAACCGGTACAGAGAAATTTTGAAGATCACTTGTGCTTGGGAACTGGAA 902
 Qy 366 ArgVal1yLysLeuLysAspTyrAlaPheIleHisPheAspG1uArgAspG1yAlaVal 385
 Db 903 CGAGTAAAGAGCTAAAGATTAATGCTTCATTCATTTGATGAGAGATGGTCTGTC 962
 Qy 386 LysAlaMetG1uGlnMetAsnG1yLysAspLeuG1uG1uLysLeuIleG1uIleValPhe 405
 Db 963 AAGGCTATGAGAAAGAAATGTAAGTCTGAGAGGAGAAATATATGAAATGTTTTT 1022
 Qy 406 AlaLysProProAspG1uLysArgLysG1uArgLysAlaGlnArgGlnAlaLysAsn 425
 Db 1023 GCTAAGCACAATCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGAGCAAGAAAT 1082
 Qy 426 GlnMetTyrAspAspTyrTyrTyrG1yProProHisMetProProProThnArgG1y 445
 Db 1083 CAATGATATATATTAATTAATTAATGCTCACTCACTATGCTCCCAACAGAGGT 1142
 Qy 446 ArgG1yArgG1yG1uArgG1yG1yTyrG1yTyrProProAspTyrTyrG1yTyrG1uAsp 465
 Db 1143 CGAGGGGGTGAAGATGAGGTGCTATGATGATCCCTCAAGATTAATTAATGATGAGAT 1202
 Qy 466 TyrTyrAspTyrTyrG1yTyrAspTyrHisAsnTyrArgG1yG1yTyrG1uAspProTyr 485
 Db 1203 TATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1262
 Qy 486 TyrG1yTyrG1uAspPheGlnValG1yAlaArgG1yArgG1yG1yArgG1yAlaArgG1y 505

Db 1263 TATGTTATGAGATTTTCAAGTTGAGCTAGAGAAAGGGGTGTAGAGAGCAAGGGGT 1322
 Qy 506 AlaAlaProSerArgG1yArgG1yAlaAlaProProArgG1yArgAlaG1yTyrSerGln 525
 Db 1323 GCTGCTCCATCCAGAGGTCTGTGGGGCTGCTCTCCCGGTGTAGAGCCGGTATTCACAG 1382
 Qy 526 ArgG1yG1yProG1ySerAlaArgG1yValArgG1yAlaArgG1yG1yLysGlnGln 545
 Db 1383 AGAGAGAGCCCTTGATTCAGCAAGAGGCGTTCGCGGCGAGAGAGGTGCCCAACAA 1442
 Qy 546 ArgG1yArgG1yG1yLysG1yValG1uAlaG1y 556
 Db 1443 AGAGCGCGGGGTAGCTGTGTGCAAGGGGTGCG 1475
 RESULT 13
 AX011753 3101 bp DNA linear PAT 06-SEP-2000
 LOCUS AX011753
 DEFINITION Sequence 156 from Patent WO955858.
 ACCESSION AX011753
 VERSION AX011753.1 GI:9998277
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 3101)
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
 Pilarsky,C.
 TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
 JOURNAL Patent: WO 955858-A 156 04-NOV-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METZGER GBS FUER GENOMFORSCHUN
 (DE); PILARSKI CHRISTIAN (DE)
 FEATURES
 source 1..3101
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 980 a 540 c 670 g 911 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,486-109 Length: 3101
 Score: 2502.00 Matches: 476
 Percent Similarity: 92.98% Conservative: 1
 Best Local Similarity: 92.79% Mismatches: 0
 Query Match: 83.90% Indels: 36
 DB: Gaps: 2
 US-09-821-687-4 (1-561) x AX011753 (1-3101)
 Qy 50 G1yLeuValAlaHisSerAspLeuAspG1uArgAlaIleGlnAlaLeuLysGlnPheAsn 69
 Db 89 GGGCTAGTGGACATAGTATGATTAAGTGAAGAGCTTATGAAGCTTTAAAGATTCAT 148
 Qy 70 G1uAspG1yAlaLeuAlaValleuGlnGlnIlePheLysAspSerAspLeuSerHisValGln 89
 Db 149 GAAGACGGCTATGGCAAGTCTTCAACAGTTTAAAGACAGAGATCTCTCAATGTTGAG 208
 Qy 90 AsnLysSerAlaPheLeuCysG1yValMetLysThnTyrArgG1uArgG1uLysGlnG1y 109
 Db 209 AACAAAAGTCCCTTTTATAGTGAAGTCATGAAGACTTACAGCAGAGAAAAACAGGG 268
 Qy 110 ThnLysValAlaAspSerSerLysG1yProAspG1uAlaLysIleLysAlaLeuLeuGln 129
 Db 269 ACCAAAGTACAGATTCCTAGTAAAGACAGATGAGGCAAAATTAAGGCACTTGGAA 328
 Qy 130 ArgThnG1yTyrThnLeuAspValThnThnG1yGlnArgLysTyrG1yG1yProPro 149
 Db 329 AGACAGGCTTACCACTTATGATGACACTGACAGAGAAAGTATGAGAACCACTTCA 388
 Qy 150 AspSerValTyrSerG1yGlnGlnProSerValG1yThnG1uIlePheValG1yLysIle 169
 Db 389 GATTCGTTATTCAGGTGAGAGGCTTCTGTGGCACTGAGATATTTGTGGGAAAGATC 448

16 AspThrThrsAlaValIleHisSerGluAsnPhenIleuLeuAspAlaGlyLeu 35
 58 GAR---ACTTCCAGTGTACTACACAGAAACATCAAGACATGATAGAGGAGGCTTC 114
 36 ProGlnIleValAlaGluIleAspGlnIleTyrValAlaGlyLeuValAlaHisSer 55
 115 CCACAGAGGCTGACAGAAAGATGATGAATTTTTCAGACAGGATGGTAGCTTATGTC 174
 56 AspLeuAsnGlnIleValAlaGluIleValAlaGluIleValAlaGluIleValAla 75
 175 GATCTTGATTAAGAGAAATCAAGTCTCAAGGAGATTATAGAGAAAGAGCTCTGTCT 234
 76 ValLeuGlnIlePheIleAspSerAspLeuSerHisValAlaGlnIleValAlaPheLeu 95
 235 GTATCTGACAGATTCACAAAGAAAGTACCTGTCTCATCTCCAGAAACAAATGCTATTTTA 294
 96 CysGlyValMetIleTyrThrTyrArgIleValGlnIleTyrIleValAlaAspSer 115
 295 TGTGAGATTATGAAGACCTACAGGCAAGAGAAACAGGCAAGAGTACAAAGAGTCC 354
 116 SerIleGlyProAspGlnIleValAlaGluIleValAlaGluIleValAlaGluIle 135
 355 ACCAAAGGCGCCATGAAGAAATCAAGCCTTGTGAGAGAGCTGCTTATACCTTTC 414
 136 AspValThrThrGlnIleValGlnIleTyrGlyProProAspSerValTyrSerGly 155
 415 GATGTACACCGGAGACAGAGAAATAGCGGCTCCACACAGAGTGTGATCTCGAGT 474
 156 GlnGlnProSerValGlyTyrGlnIlePheValGlyIleProAspLeuPheGln 175
 475 GTCCAGCCTGAGTATGGGACAGAGTCTTGTAGTAAATCCCAAGAGCTGTATGAG 534
 176 AspGlnIleValIleProLeuPheGlnIleValAlaGlyProIleThrAspLeuMetMet 195
 535 GATGAATATAGTCCCTTTTGAAGAGCTGTGCTATCTGGAGATCTACCGCTCATGATG 594
 196 AspProLeuThrGlnIleValAsnAsnGlyTyrValAlaPheValThrPheCysThrIleValAla 215
 595 GATCCATCTCTGTGTGAGAAACAGAGCTATGCTATCTATCTCTGTGAAAGAAAGCT 654
 216 AlaGlnIleValAlaValIleValIleTyrAsnAsnIleGlnIleAspSerGlyIleValAla 235
 655 GCCACAGAAACAGATTAACTGTGTGACAGATGATGATGCTGCTGTAACACCTGGGA 714
 236 ValCysIleSerValAlaAsnAsnArgLeuPheValGlySerIleProIleValSerIleThr 255
 715 GTGTGATTTCTGTGCAACACAGACTTTTGTGTGATCTTCCAAAGAAATTAACACT 774
 256 LysGlnIleIleLeuGlnIlePheSerIleValThrGlnIleLeuThrAspValIleLeu 275
 775 AAGGAAACACTTGTGAAAGCTTCAAGAACTCAAGAGGATTTGTGAGACGTTATTTCTC 834
 276 TyrHisGlnProAspArgIleValIleValAsnArgIlePheCysPheLeuGlnIleValAsp 295
 835 TATCATCAACCCGATGCAAAAGAAAGATCGGGGCTTCTCTTGAATATGAGAT 894
 296 HisIleThrAlaIleGlnIleArgIleArgLeuMetSerGlyIleValIleValIleTyrGly 315
 895 CACAAGTCAGACACAAAGCAGACGCGCTGATGAGTGAAGAAAGTCAAGAGTGGGGA 954
 316 AsnValGlyThrValGlnIleValAspProIleGlnIleAspProAspProGlnIleValAla 335
 955 AATGTGTTCAGTAAAGTGGCTGACCCCGTGGAGGAGCGGATCCAGAAAGTCAAGGCA 1014
 336 LysValIleValIlePheValIleArgAsnIleValAsnThrValThrGlnIleIleLeuGln 355
 1015 AAGGTAAAGATTATTTGTGAGAACTTGGCAACTACGCTGACAGAAAGAAATTTAGAG 1074
 356 LysSerPheSerIlePheGlyIleValIleValIleValIleValIleValIleValIleVal 375
 1075 AAGTCAATTTCTAGTTGAGAAACCTGAGCGAGTGAAGAAAGTTGAAGATTTATGCGTTC 1134

376 IleHisPheAspGlnIleValAlaValIleValAlaMetGlnIleValAsnGlyIleValAsp 395
 1135 GTTCAATTTGAGACAGGAGAGCTGCTTTAGCGTATGATGAAATGAAATGAAAGAAAGAA 1194
 396 LeuGlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 415
 1195 ATGAAAGGGAAGAAATTTGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 1254
 416 ArgValAlaGlnIleValIleValIleValIleValIleValIleValIleValIleVal 435
 1255 CGCAGCTGTGACAGAGCTCCCAAGACCTGCGTATGAAGATTTATATATACACCT 1314
 436 ProProHisMetProProProProProProProProProProProProProProProPro 455
 1315 CTTCCCGGATGCTCCTCATGAGAGCCCAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1374
 456 TyrProProAspTyrTyrGlyIleValIleValIleValIleValIleValIleValIle 474
 1375 TACCTCCAGATTTATATATGCTATGAGAACTATGATGATGATGATGATGATGATGAT 1434
 475 HisAsnTyrArgGlyIleValIleValIleValIleValIleValIleValIleValIle 493
 1435 CACGACTATCGTGAAGCTATGAAGATCCCTACTACGCTATGATGATGATGATGATGAT 1494
 494 GlyAlaArgGlyValGlyIleValIleValIleValIleValIleValIleValIleVal 512
 1495 AGAGAAAGAGCA---GAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1551
 513 GlyAlaIleProProArgIleValAlaGlyTyrSerGlnIleValIleValIleValIle 531
 1552 GAGGACACACCTCCAAAGAGTATGAGCTGCTATTCACAGAGGAGGAGCACTTTGGAGCA 1611
 532 AlaArgIleValIleValIleValIleValIleValIleValIleValIleValIleVal 550
 1612 CCAAGAGGCTCAAGAGGCTGAGAGAGGAGGCTCTGACACACAGAGAGAGCGGTGCT 1671
 551 LysGly 552
 1672 CGTGGA 1677

RESULT 15
 AF000364
 LOCUS
 DEFINITION
 Homo sapiens heterogeneous nuclear ribonucleoprotein R mRNA,
 complete cds.
 ACCESSION
 AF000364
 VERSION
 AF000364.1 GI:2697102
 SOURCE
 ORGANISM
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2644)
 Hassefeld,W., Chan,B.K., Mathison,D.A., Portman,D., Dreyfuss,G.,
 Steiner,G. and Tan,E.M.
 Molecular definition of heterogeneous nuclear ribonucleoprotein R
 (hnRNP R) using autoimmune antibody: immunological relationship
 with hnRNP P
 Nucleic Acids Res. 26 (2), 439-445 (1998)
 JOURNAL
 MEDLINE
 PUBMED
 9421497
 98083170
 REFERENCE
 2 (bases 1 to 2644)
 Hassefeld,W., Chan,B.K.I. and Tan,E.M.
 Direct Submission
 Submitted (18-Apr-1997) Molecular and Experimental Medicine, The
 Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla,
 CA 92037, USA
 FEATURES
 source
 1..2644
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="BL14"
 /cell_line="T24"

CDS

62..1963
/note="hmbNP R"
/codon_start=1
/product="heterogeneous nuclear ribonucleoprotein R"
/protein_id="AAC39540.1"
/db_xref="GI:2697103"

translation="MANOVGNVAVOLKEEPEMDTSVTHTEHYKTLIEAGLPORVAB
RLDEIFOTGLVAVVDLDERALIDALRENEEGALSVICOFKESPLSHVONKSAFLCGVM
KTYRREKQSSKVOEESKQDEKAKIRALLERTGYTIDYTGQRKGGPPDPSYISVQ
PGIGIEYFVGKIPFDLYDELVPLFEFAGITWDRLMDPLSGONKGFATITCGKEA
AOEAVKLCDSYEIIRPGHILGVCISVANNRLIVGSIIPNKTKENILIEFSKVTGLVDY
ILYHOPDKKKKNGPCFLYEYDEKSAQAARRLMSGKVKWGVAVVEMADPYEEDP
EWAKYVLFVRLNATVTEBEILEKSPSEFSGKLEKDYAFVHEDRGAIVKAMD
ENMKIEGEEIEIVLAKPPDKRKERKARASRPAVEDYVYHPPEMPPIRRC
RGGRGSGYGPDPYGYEDYDYDYGTDHYDHYGKGYEDYGYDDGYATVRGGGGG
RGAPPPRGAPPRGAGYSQRPGLPGRSGRGSGGPAQOQKRGSRGSRNG
GNVGKRRKADGYNQPSKRRQTNNOQWGSQPIAQQPLQGGGYSQNGYNNDOERY
QDYYGQOWK"

BASE COUNT 815 a 476 c 665 g 688 t

ORIGIN

Alignment Scores:
Pred. No.: 3,08e-104 Length: 2644
Score: 2403.00 Matches: 458
Percent Similarity: 90.23% Conservative: 50
Best Local Similarity: 81.35% Mismatches: 41
Query Match: 80.58% Indels: 14
DB: Gaps: 10

US-09-821-687-4 (1-561) x AF000364 (1-2644)

QY 1 MetAlaThrGluHISValAsnGlyAsnGlyThr-----GluGluProMet 15
DB 62 ATGGCTATTCG---GTGAATGTGTAATCGGTACAGTTAAAGAGAGAGAAACCAATG 118
QY 16 AspThrThrSerAlaValIleHisSerGluAsnPhenGlnThrLeuAspAlaGlyLeu 35
DB 119 GAT---ACTTCAGTGTACACACAGACACATCAAGACACATGATAGAGCGGCTC 175
QY 36 ProGlnLysValAlaGluLysLeuAspGluIleTyrValAlaGlyLeuValAlaHisSer 55
DB 176 CCACAGAGGTGGCAGAAAGCTGATCAATAATTTTCAGACAGATTTGGTCACTTATGTC 235
QY 56 AspLeuAspGluArgAlaIleGluAlaLeuLysGluPheAsnGluAspGlyAlaLeuAla 75
DB 236 GATCTTGAATGAAGACATTAATGCTCTCAGGAAATTAATGAAGAGAGCTGCTCT 295
QY 76 ValLeuGlnGlnPheLysAspSerAspLeuSerHisValGlnAsnLysSerAlaPheLeu 95
DB 296 GTACTACAGCAGTTCAAGGAAGTGACTTATCAATGTTCAAGAACAAAGTCATTTTA 355
QY 96 CysGlyValMetLysThrTyrArgGlnArgGluLysGlnGlyThrLysValAlaAspSer 115
DB 356 TGTGAGATTAAAGAACCTTACAGCAGAGAGAAACAGGAGAGAGTCCAGAGCTCC 415
QY 116 SerLysGlyProAspGluAlaLysIleLysValAlaLeuLysGlnArgThrGlyTyrThrLeu 135
DB 416 ACAAGGAGCCTGAATGAAGCAAGTCAAGGCTTGCTTGAAGAGACGTTATTACTTG 475
QY 136 AspValThrThrGlnArgLysLysTyrGlyProProAspSerValTyrSerGly 155
DB 476 GATGTAACCAAGAGCAGAGAAATATGTGTCTCTCACACAGACAGTGTATCTGGC 535
QY 156 GlnGlnProSerValGlyThrGluIlePheValGlyLysIleProArgAspLeuPheGlu 175
DB 536 GTGCAACCTGGAATGGAAACGAGATATTGTAGCAAAATACCAAGAGATTATATAG 595
QY 176 AspGluLeuValProLeuPheGluLysAlaGlyProIleTyrAspLeuArgLeuMetMet 195
DB 596 GATAGATTGGTCCCTTTTGAAGAGCCGAGCCCATTTGGAGATCAAGCTTATATGATG 655
QY 196 AspProLeuThrGlyLeuAsnArgGlyTyrAlaPheValThrPheCysThrLysGluAla 215

DB 656 GATCACTGTCCGGTCAAGATAGAGGGTATGATTCATCACTTCTGTGGAAAGAGAGCT 715
QY 216 AlGlnGluAlaValLysLeuTyrAsnAsnHisGluIleArgSerGlyLysHisIleGly 235
DB 716 GCACAGGAACCTCGAAGACGTGTGACAGGTATGAATAATCGCCCTGTAAACCTTGG 775
QY 236 ValCysIleSerValAlaAsnAsnArgLeuPheValGlySerIleProLysSerLysThr 255
DB 776 GTGTGATTTCTGTGGCAACACAGACTTTTGTGGATTCATTTCCAGAAATTAAGACT 835
QY 256 LysGlnGlnIleLeuGlnGluPheSerLysValThrGluLysLeuThrAspValIleLeu 275
DB 836 AAGAAACATTTTGGAAAGATTCAGTAAAGTCACAGAGGTTTGGTGACGTTATTC 895
QY 276 TyrHisGlnProAspAspLysLysLysAsnArgGlyPheCysPheLeuGluTyrGluAsp 295
DB 896 TATCATCAACCCCATGACAAAGAAAGAAATCGGGGGTTCCTGCTTCAATATYGAGAT 955
QY 296 HisLysThrAlaAlaGlnAlaArgArgLysLeuMetSerGlyLysValLysValTyrGly 315
DB 956 CACAGTCAGCAGACACAGCCAGACCCCGCTGATAGTGAAAGTAAAGTGTGGGA 1015
QY 316 AsnValGlyThrValGluTrrAlaAspProIleGluAspProAspProGluValMetAla 335
DB 1016 AATGTAGTTACAGTTGAATGCGCTGACCTGTGGAAGAACACAGATCAGAGTCATGGCT 1075
QY 336 LysValLysValLeuPheValArgAsnLeuAlaAsnThrValThrGluGluIleLeuGlu 355
DB 1076 AAGGTAAAGTTTGTGTGGAAACCTGGCTACACGCTGACAGAGAAATATGGA 1135
QY 356 LysSerPheSerGlnPheGlyLysLeuGluArgValLysLysLeuLysAspTyrAlaPhe 375
DB 1136 AAGTCATTTTCTGAATTTGGAATTCGAAAGGTAAAGTGAAGTTGAAGTTATGATTT 1195
QY 376 IleHisPheAspGluLysArgAspGlyAlaValLysAlaMetGlnGluLysAsnGlyLysAsp 395
DB 1196 GTTCATTTTGAAGACAGAGAGACAGCGTGTAAAGCTATGATGATGAATGATGACAAAGA 1255
QY 396 LeuGlnGlyGluAsnIleGluIleValPheAlaLysProProAspGluLysValGly 415
DB 1256 ATGAAGGGGAAATAATGAATATGCTTAAAGCCACACACAGAAAGAAAGAG 1315
QY 416 ArgLysLysGlnArgGlnAlaAlaLysAsnGlnMetTyrAspAspTyrTyrTyrGly 435
DB 1316 CGCCAAAGCTCTGAAGAGGCTTCCAGAGAGACCTGCTATGAATATTAATCAACCT 1375
QY 436 ProProHisMetProProProThrArgGlyArgGlyArg---GlyLysArgGlyGlyTyr 454
DB 1376 CTTCTCGCATGCACTCCAAATTAGAGGTGGGGTCTGTGGGGGAGAGGTGATAT 1435
QY 455 GlyTyrProProAspTyrTyrGlyTyrGluAspTyrTyr---AspTyrTyrGlyTyrAsp 473
DB 1436 GGCTAACCTTCAAGATTAACAGCTATGAAGATTAATGATGATTAATGATGAT 1495
QY 474 TyrHisAsnTyrArgGlyGlyTyrGluAspProTyrTyrGlyTyrGluAsp---PheGln 492
DB 1496 TATCACACATATGTGAGGCTATGAAGATCCTTACTAGGCTATATATATGCTATGCA 1555
QY 493 ValGlyAlaArgGlyArgGlyArgGlyAlaArgGlyAla---AlaProSerArgGly 511
DB 1556 GTAAGAGGAAGAGA---GAGAGAGGGGAGGGGAGGAGTCTCACACACCAAGAGGGG 1612
QY 512 ArgGlyAlaAlaProProArgGlyArgAlaGlyTyrSerGlnArgGlyPro---Gly 530
DB 1613 AGGGAGACACACCTCAAGAGGTAGAGCTGCTATATCAAGAGGGGGGACCTTTGGGA 1672
QY 531 SerAlaArgGlyValAlaArgGlyAlaArgGlyGly---AlaGlnGlnGlnArgGlyArgGly 549
DB 1673 CCACCAAGAGCTCTTAGGGGTGGCAGAGGGGGTCTGCTCAACAGCAGAGGCGCGTGT 1732
QY 550 GlyLysGly 552
DB 1733 TCCCGTGA 1741

Mon Apr 28 06:48:47 2003

Search completed: April 28, 2003, 06:50:02
Job time : 8634 secs

us-09-821-687-4.rge

Page 21

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2003, 16:49:41 ; Search time 17371.6 Seconds (with about 34 comments)

5783.170 Million cell updates/sec

Title: US-09-821-687-3
 Perfect score: 2453

Sequence: 1 gcgggttctgcacagcgc.....ggtaccgccggtacagagct 3452

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: gub_ba:*

```
3: gub_in: *
4: gub_out: *
```

*_ov:gb

7: gbo_ph:*

```

9:  gb_pr: *
10:  gb_pr: *

```

```
11: gjb_bts:*
```

```
13: gfb_un: *
14: gfb_un: *
```

15: em_ba: *

17: em_hum: 4

```
19: em_mu:*
```

```
21: em_or: *
```

23: em_pat: *
24: em_ph: *25: em_pl: *
26: em_ro: *

27: em_bts: *

```
29: em_vi: *
30: em_htr: *
```

31: em_htg_1
32: em_htg_2

```
33: em_htg_n
34: em_htg_n
```

35: em_htg_1
36: em_htg_137: em_htg_v
38: em_gv.*

39: em_htgo-
40: em_htgo-

41: em_hfgo_

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	3452	100.0	3452	10	AB035725	AB035725 Mus muscu	
2	2741.4	79.4	3319	9	BC032843	BC032843 Homo sapi	
3	2542.2	73.6	3101	6	AX011753	AX011753 Sequence	
4	1953.8	56.6	2221	9	AF155568	AF155568 Homo sapi	
5	1886.6	54.7	2195	10	BC004001	BC004001 Mus muscu	
6	1826.2	52.9	2079	6	AB177851	AB177851 Sequence	
7	1802	52.2	103819	9	HS3775C13	HS3775C13 Homo sapi	
8	1802	52.2	212859	2	ALJ90737	ALJ90737 Homo sapi	
9	1674	48.5	2532	9	AF037448	AF037448 Homo sapi	
10	1608	46.6	2373	10	AF093821	AF093821 Mus muscu	
11	1584.8	45.9	1686	9	AY034483	AY034483 Homo sapi	
12	1552.2	45.0	1872	9	AY034481	AY034481 Homo sapi	
13	1456.2	42.2	2232	10	AF408431	AF408431 Mus muscu	
14	1356.4	39.3	74337	9	ALJ36082	ALJ36082 Human DNA	
15	1343.4	38.9	1767	9	AY034482	AY034482 Homo sapi	
16	1156	33.5	221309	2	AC122217	AC122217 Mus muscu	
17	1033	29.9	175221	2	AC111832	AC111832 Rattus no	
18	957.6	27.7	151297	2	AC118904	AC118904 Rattus no	
19	896	26.0	1899	10	AF441128	AF441128 Mus muscu	
20	875	25.3	2644	9	AF000364	AF000364 Homo sapi	
21	856	24.8	2722	9	BC001449	BC001449 Homo sapi	
22	853.8	24.7	2471	10	BC004679	BC004679 Mus muscu	
23	816	23.6	181810	2	AC126148	AC126148 Rattus no	
24	762.4	22.1	121304	2	AC079134	AC079134 Mus muscu	
25	762.4	22.1	196488	2	AC084410	AC084410 Mus muscu	
26	742.2	21.5	202374	10	AC122868	AC122868 Mus muscu	
27	736.4	21.3	934	6	AX013715	AX013715 Sequence	
28	636.8	18.4	139996	9	AL161799	AL161799 Human DNA	
29	570.6	16.5	129373	2	AC127932	AC127932 Rattus no	
30	484	14.0	110783	2	AC129693	AC129693 Rattus no	
31	457.2	13.2	211030	2	AC125524	AC125524 Mus muscu	
32	426	12.3	118429	2	AC119124	AC119124 Rattus no	
33	398.4	11.5	57824	2	AC101053	AC101053 Mus muscu	
34	381.8	11.1	2606	3	AY058477	AY058477 Drosophil	
35	368	10.7	384	6	AR079026	AR079026 Sequence	
36	309.2	9.0	413	6	AR079027	AR079027 Homo sapi	
37	289.4	8.4	2467	9	BC028588	BC028588 Homo sapi	
38	250.8	7.3	1785	10	AY028945	AY028945 Rattus no	
39	250.8	7.3	1863	10	AY029098	AY029098 Rattus no	
40	247	7.2	175221	2	AC111832	AC111832 Rattus no	
41	242.4	7.0	1770	9	HS4272079	HS4272079 Homo sapi	
42	242.4	7.0	1794	9	HS4272078	HS4272078 Homo sapi	
43	242.4	7.0	1945	9	AF0209192	AF0209192 Homo sapi	
44	240.8	7.0	1761	9	AF0271789	AF0271789 Homo sapi	
45	233	6.7	1911	9	AF0271790	AF0271790 Homo sapi	

ALIGNMENTS

RESULT		1
LOCUS	AB035725	
DEFINITION	AB035725	3452 bp mRNA linear ROD 31-MAR-2006
ACCESSION	Mus musculus SYNCRIP mRNA, complete cds.	
VERSION	AB035725	
KEYWORDS	AB035725..1 GI:6576814	
SOURCE	SYNCRIP.	
ORGANISM	Mus musculus (strain:DY) postnatal 21 day cDNA to mRNA.	
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (sites)	
AUTHORS	Mizutani A., Fukuda M., Iwata K., Shiraiishi Y. and Mukoishiba K.	
TITLE	SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptoagmin	

QY 1561 TANGATTATGATTAACCACTAATCGTGTGATATGAAGATCACTATGTTATGAA 1620
 DB 1561 TANGATTATGATTAACCACTAATCGTGTGATATGAAGATCACTATGTTATGAA 1620
 QY 1621 GATTTTCAAGTTGAGCTAGAGAGAGGGGTGTAGAGAGCAAGGGGTGCTCCATCC 1680
 DB 1621 GATTTTCAAGTTGAGCTAGAGAGAGGGGTGTAGAGAGCAAGGGGTGCTCCATCC 1680
 QY 1681 AGAGGTCTGAGGGCTGCTCCCGGTGTAGAGCCGGTTATTCAGAGAGAGGGCCCT 1740
 DB 1681 AGAGGTCTGAGGGCTGCTCCCGGTGTAGAGCCGGTTATTCAGAGAGAGGGCCCT 1740
 QY 1741 GGATTCAGAGAGAGGGCTGCTCCCGGTGTAGAGAGAGGGCCCAACAAGAGGGCCGAG 1800
 DB 1741 GGATTCAGAGAGAGGGCTGCTCCCGGTGTAGAGAGAGGGCCCAACAAGAGGGCCGAG 1800
 QY 1801 GGAAGAGGGGTGAGAGGGGTGCTGACCTGTTCATGAAGAGCTGATTTGTGGG 1860
 DB 1801 GGAAGAGGGGTGAGAGGGGTGCTGACCTGTTCATGAAGAGCTGATTTGTGGG 1860
 QY 1861 ATTACACCAAGAGCTGAG 1920
 DB 1861 ATTACACCAAGAGCTGAG 1920
 QY 1921 CGGATGTATGAG 1980
 DB 1921 CGGATGTATGAG 1980
 QY 1981 TACGGAGAGATGAG 2040
 DB 1981 TACGGAGAGATGAG 2040
 QY 2041 AGGCAAGATTAATTAAG 2100
 DB 2041 AGGCAAGATTAATTAAG 2100
 QY 2101 ATAATCTTCAAG 2160
 DB 2101 ATAATCTTCAAG 2160
 QY 2161 TAAGTAGATTTTCCCTCTTTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
 DB 2161 TAAGTAGATTTTCCCTCTTTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
 QY 2221 CTCTTGAATTTGAG 2280
 DB 2221 CTCTTGAATTTGAG 2280
 QY 2281 TAAATTTTATTTCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 DB 2281 TAAATTTTATTTCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 QY 2341 TCACTTTGAATTTGAG 2400
 DB 2341 TCACTTTGAATTTGAG 2400
 QY 2401 GCATTCAGATAGTGTCTTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
 DB 2401 GCATTCAGATAGTGTCTTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
 QY 2461 GCAGAGAGATGAG 2520
 DB 2461 GCAGAGAGATGAG 2520
 QY 2521 AAGGCTTGTAAATATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
 DB 2521 AAGGCTTGTAAATATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
 QY 2581 TTGCTTTTATTTCAAG 2640
 DB 2581 TTGCTTTTATTTCAAG 2640

QY 2641 TTGCTTTTATTTGAG 2700
 DB 2641 TTGCTTTTATTTGAG 2700
 QY 2701 TGTATATATGTTTAG 2760
 DB 2701 TGTATATATGTTTAG 2760
 QY 2761 CATCCGAATTTGATCAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 DB 2761 CATCCGAATTTGATCAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 QY 2821 AATTGACAG 2880
 DB 2821 AATTGACAG 2880
 QY 2881 TAGTATATGAG 2940
 DB 2881 TAGTATATGAG 2940
 QY 2941 ATTGCTGATCACTATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
 DB 2941 ATTGCTGATCACTATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
 QY 3001 TTGTAGATATGAG 3060
 DB 3001 TTGTAGATATGAG 3060
 QY 3061 CATGAGGTGAG 3120
 DB 3061 CATGAGGTGAG 3120
 QY 3121 AAGTCTTAAGTTGATAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 DB 3121 AAGTCTTAAGTTGATAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 QY 3181 TAAAG 3240
 DB 3181 TAAAG 3240
 QY 3241 GCAACAGATGAG 3300
 DB 3241 GCAACAGATGAG 3300
 QY 3301 TAAATGCAATTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
 DB 3301 TAAATGCAATTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
 QY 3361 AATCATTTTGAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
 DB 3361 AATCATTTTGAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
 QY 3421 AACTGCTAATGAG 3480
 DB 3421 AACTGCTAATGAG 3480

RESULT 2
 BC032643
 LOCUS BC032643 3319 bp mRNA linear PRI 27-JUN-2002
 DEFINITION Homo sapiens, similar to NS1-associated protein 1, clone MGC:45213
 IMAGE:5495201, mRNA, complete cds.
 ACCESSION BC032643
 VERSION BC032643.1 GI:21619167
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3319)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian

QY	2455	TCAGTAGCAGAGAGATGAACTATACCCCTCTGTAAGTACATAGAAAATCTTTGAGAT	2514
Db	2296	TCAGTAGCAGAGAGATGAACTATACCCCTCTGTAAGTACATAGAAAATCTTTGAGAT	2355
QY	2515	AAAGCAGAGCGCTTGTTAAATATGATATGAGGGTAAAGATTTTAAATATACCAATGTACAT	2574
Db	2356	AAAGCAGAGCGCTTGTTAAATATGATATGAGGGTAAAGATTTTAAATATACCAATGTACAT	2415
QY	2575	TCCTAGTGGCCTTATAGTTTACAGAGGCTGTAGACTTCTCTCATGACCATCATPACAGGCC	2634
Db	2416	TCCTAGTGGCCTTATAGTTTACAGAGGCTGTAGACTTCTCTCATGACCATCATPACAGGCC	2475
QY	2635	TTGCTTTTGTGCTATTTTGTGTGCTGAAAAAGACGCTTGCTTCTTCAGATATTTAGTATTA	2694
Db	2476	TTGCTTTTGTGCTATTTTGTGTGCTGAAAAAGACGCTTGCTTCTTCAGATATTTAGTATTA	2535
QY	2695	TTTGGATGTATTAATAGTTTAGCAAGATGTACTTTTGTATAGACATCAGATGTTCAAAAA	2754
Db	2536	TTTGGATGTATTAATAGTTTAGCAAGATGTACTTTTGTATAGACATCAGATGTTTC	2592
QY	2755	AAAGTCATCCGACATTTGTATCTAATAATCTGACAGTGTCCCTTATATAAAAGTCAGACTTAA	2814
Db	2593	AAAGTCATCCGACATTTGTATCTAATAATCTGACAGTGTCCCTTATATAAAAGTCAGACTTAA	2652
QY	2815	ACTGCAATTTGTACACGCAAAAGCTTCAATTTGGATATTTTGAAGTTTTCATTAATCAT	2874
Db	2653	ACTGCAATTTGTACACGCAAAAGCTTCAATTTGGATATTTTGAAGTTTTCATTAATCAT	2712
QY	2875	AGAATTTAGTATATAGGCTGTAGTTAGCTTTTAGGTAAAGGTATGTTCAATAGTGA	2934
Db	2713	AGAATTTAGTATATAGGCTGTAGTTAGCTTTTAGGTAAAGGTATGTTCAATAGTGA	2772
QY	2935	TTTGTATTGTGCTGATCACTATTAATAATGTGATTCAGCTTCCATTTCTATGACAGTCAAT	2994
Db	2773	TTTGTATTGTGCTGATCACTATTAATAATGTGATTCAGCTTCCATTTCTATGACAGTCAAT	2832
QY	2995	GATACTTGTAGAAATAGAGTACATCATTTGTGCTATGTTTAAATTTCTTAAAGCACTT	3054
Db	2833	GATACTTGTGTAGAGTATAGTATCAATCATTTGTGCTATGTTTAAATTTCTTAAAGCACTT	2892
QY	3055	TGATACACATGAGTGTTCAGTGTGTAGAGCATCTCTATTTGATTCACCCCAAAAAATTTT	3114
Db	2893	TGATACACATGAGTGTTCAGTGTGTAGAGCATCTCTATTTGATTCACCAACCTTC	2951
QY	3115	TTTGCAGATCCCTAAGTTGATATAGCTTAAAGTCAAAAGT	3173
Db	2952	TTTGCAGATCCCTAAGTTGATATAGTAAAGTAAAGTAAATATATAGTTTCAATTAGGA	3011
QY	3174	CTTGCTGTAAAGAAACAACCTT	3232
Db	3012	CTTGCTGTAAAGAAATCCCTTCCCTTCCCTTCCCAAAAGGATATCTGACATTAATACATA	3071
QY	3233	CCCAATAGGCAACAAGATGAGATCAGAGC	3291
Db	3072	CCCAATAGGCAACAAGATGAGATCAGAGCCTTATCTTAATTAAGTTTATACACACA	3131
QY	3292	GTTTCCCCAGTAAATGCAAAATTTTACAGAAAAATTTAGACATGTCATATGTTCAAAAATGCTC	3351
Db	3132	GTTTCCCCAGTAAATGCAAAATTTTACAGAAAAATTCAGACATGTCATATGTTCAAAAATGCTC	3191
QY	3352	ATGGCAAAATCATATTTTGCATTCCTGTGCAAAATTAATGTTTATATAC	3398
Db	3192	ATGGCAAAATCATATTTTGCATTCCTGTGCAAAATTAATGTTTATATAC	3238

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 3101) Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.
TITLE	Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL	Patent: WO 995858-A 156 04-NOV-1999; SCHMITT AMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)
FEATURES	location/Qualifiers
source	1..3101 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	980 a 540 c 670 g 911 t
ORIGIN	
Query Match	73.6%; Score 2542.2; DB 6; Length 3101;
Best Local Similarity	92.8%; Pred. No. 0;
Matches 2890; Conservative	0; Mismatches 93; Indels 132; Gaps 16;
Db	238 GCAGGCGTAGTGTGACATATGATGATTAGATGAAAGCATTCGAAGCTTTAAAGAGTTC 357
Db	86 GGAAGGCGTAGTGTGACATATGATGATTAGATGAAAGCATTCGAAGCTTTAAAGAGTTC 145
Db	358 AATGAAGCGCGCATTTGGCAGTGTCTTCAACGTTTAAAGACAGTGATCTCTCATGTT 417
Db	146 AATGAAGCGCGCATTTGGCAGTGTCTTCAACGTTTAAAGACAGTGATCTCTCATGTT 205
Db	418 CAGAACAAAGCGCTTTTATGTGAGCAGAGAAGACTTACAGGACAGAGAAAGACG 477
Db	206 CAGAACAAAGCGCTTTTATGTGAGCAGAGAAGACTTACAGGACAGAGAAAGACG 265
Db	478 GGGACCAAAGTAGCAGACTCTAGTAAAGACAGATGAGGCAAGATTAGGACCTTTTG 537
Db	266 GGGACCAAAGTAGCAGACTCTAGTAAAGACAGATGAGGCAAGATTAGGACCTTTTG 325
Db	538 GAAAGAACGCGTACACACTGATGTGACTACAGGTACAGGAGATGAGGACCACT 597
Db	326 GAAAGAACGCGTACACACTGATGTGACTACAGGAGATGAGGACCACT 385
Db	598 CAGAGATCCGTTTATTCAGGTACAGGACCTTCTGTGGACATGATATTTGGGGAG 657
Db	386 CAGAGATCCGTTTATTCAGGTACAGGACCTTCTGTGGACATGATATTTGGGGAG 445
Db	658 ATCCCGAGAGATCTGTGTGAGTAGGCTGTTCCTATTTAGAGAACTGACCTATA 717
Db	446 ATCCCGAGAGATCTGTGTGAGTAGGCTGTTCCTATTTAGAGAACTGACCTATA 505
Db	718 TGGAGATCTTGGTTTATGATGATCGCGCTCACTGATCTCAACAGAGATTATGCGTTGTC 777
Db	506 TGGAGATCTTGGTTTATGATGATCGCGCTCACTGATCTCAACAGAGATTATGCGTTGTC 565
Db	778 ACTTTTGTGACAAAGAACAGACAGAGGCGTGTAACTGTATATATCATGAAT 837
Db	566 ACTTTTGTGACAAAGAACAGAGGCGTGTAACTGTATATATCATGAAT 625
Db	838 CGTCCGGGAGAGCAATGAGTGTGATCTGATCTGATGCAACATATGAGCTTTTGAGGAC 897
Db	626 CGTCCGGGAGAGCAATGAGTGTGATCTGATCTGATGCAACATATGAGCTTTTGAGGAC 685
Db	898 TCGATTCCTAAGAGTAAACCAAGAGACAGATCTTGAGAGATTTAGCAAGTGCAGAG 957
Db	686 TCGATTCCTAAGAGTAAACCAAGAGACAGATCTTGAGAGATTTAGCAAGTGCAGAG 745
Db	958 GGTCTCAAGATGCTATTTTATACCAACCTGATGACAGAGAAAGAAACAGAGGCTTT 1017
Db	746 GGTCTCAAGATGCTATTTTATACCAACCTGATGACAGAGAAAGAAACAGAGGCTTT 805
Db	1018 TGTCTTCTGAATATGAAATGACAAACAGCTGCCAGGCAAGAGTACGTAAGAT 1077
Db	805 TGTCTTCTGAATATGAAATGACAAACAGCTGCCAGGCAAGAGTACGTAAGAT 842

Qy	1078	GGTAAAGTCAAAGCTCGGGGAAATGTGGAACTGTGAGTGGGCTGATCTATTAAGAT	1137
Db	843	-----	842
Qy	1138	CTCGATCTGAAATTATGCGAAAGTAAAGTGTCTGTTGTACGCAACCTTGCAACG	1197
Db	843	-----AGGTAAAGGCTGTTGTACGCAACCTTGCAACACT	880
Qy	1198	GTAACAGAAATTTTAGAAAAGTCATTATGCTTGAGAAATCGAAGCGATTAAG	1257
Db	881	GTAACAGAAAGATTTTAGAAAAGCATTTAGTCAGTTTGGGAAATCGAAGCGATGAG	940
Qy	1258	AAGCTAAAGATTATGCTTTCATTCATTTGATGAGAGATGTGTCTGTCAAGGCTATG	1317
Db	941	AAGTTAAAGATTTATGCTGTTCAATTCATTTGATGAGCAGATGTGCTGTCAAGGCTATG	1000
Qy	1318	GAAGAAATGAATGTGAAGACCTTGAGAGAGAAATATTGAATATGTTTTGCTAACCA	1377
Db	1001	GAAGAAATGAATGTGCAAGACTTGAGAGAGAAATATTGAATATGTTTTGCAAGCA	1060
Qy	1378	CCAGATCAGAAAGAGAAAGAAAGAAAGCTCAGAGCAGACGCAAGATCAATGTAT	1437
Db	1061	CCAGATCAGAAAGAGAAAGAAAGAAAGCTCAGAGCAGACGCAAGAAATCAATGTAT	1120
Qy	1438	GATGATTACTACTATTTATGCTCACCCTCATATGCTCTCCCAACAGAGTGCAGGCGT	1497
Db	1121	GACATTTACTACTATTTATGCTCACCCTCATATGCTCTCCCAACAGAGTGCAGGCGT	1180
Qy	1498	GGAGGTGAGAGTGGCTATGGAATACCTCCAGATTTATTTGATACGAAGATTAATGAT	1557
Db	1181	GGAGGTGAGAGTGGTATGGAATACCTCCAGATTTATTTGATGAAAGATTAATGAT	1240
Qy	1558	TATTTATGTTATGTTATACATACTATCGTGTGATATGAAGATCCATCTATGTTAT	1617
Db	1241	TATTTATGTTATGTTATACATACTATCGTGTGATATGAAGATCCATCTATGTTAT	1300
Qy	1618	GAAATTTTCAAGTTGAGCTAGAGAAAGGGTGTGTAAGAGCAAGGGGTGCTGCTCA	1677
Db	1301	GAAATTTTCAAGTTGAGCTAGAGAAAGGGTGTGTAAGAGCAAGGGGTGCTGCTCA	1360
Qy	1678	TCCAGAGTGTGGGGCTGCTCTCCCGTGTAGAGCCGGTTATTCACAGAGAGAGGC	1737
Db	1361	TCCAGAGTGTGGGGCTGCTCTCCCGCGGTATGAGCCGGTTATTCACAGAGAGAGT	1420
Qy	1738	CTTGATTCAGAAAGGCGTTGCGGGTGCAGAGAGGCTGCCAACAACAAGGCGCG	1797
Db	1421	CTTGATTCACAAGAGGCGTTGCAAGGTGTGAGAGAGGCTGCCAACAACAAGGCGCG	1480
Qy	1798	GG-----GGGAAAGAGGGGTGCAGAGCCGGTCTGACCTGTTACAATGAACTGACTTGAT	1854
Db	1481	GGGCAAGGAAAGGGGTGCAGAGCCGGTCTGACCTGTTACAATGAACTGACTTGCTA-	1539
Qy	1855	TGTGGATTACACAGAGCTTGCAGTGAGATGTGTAAGGAAATCAGCAACTTAA	1914
Db	1540	TGTGGATTACACAGAGCTTGCAGTGAGATGTGTAAGG- AATCAGCAACTTAA	1598
Qy	1915	ATATCTCGAGATGTAAGAGCAATTTCTATGCAAGACCTCTTAAGAGATCATGGA	1974
Db	1599	ATATGTCGGCTGTAAGAGCAATTTCTATGCAAGACCTTCTTAAGAGATCATGGA	1658
Qy	1975	ATCAAAATACGGGACATTTGAACTTAATCTTGAACTTGTAATTTCTTTAACAATTT	2034
Db	1659	ATCAAAATACGGGACATTTGAACTTAATCTTGAACTTGTAATTTCTTTAACAATTT	1718
Qy	2035	CTTGCAAGTCAGTTATTTAACTAAAGCTACTCTATTTTCCAAATGTGTTCAA- AAA	2093
Db	1719	CTTGCAAGTCAGTTATTTAACTAAAGCTACTCTATTTTCCAAATGTGTTCAAAGAA	1778
Qy	2094	ATTCCTATTAATCTTCAGAGATGTATCTTAATTAAGAAATTAAGTTGTCTCTTTAAAAA	2153
Db	1779	ATTCCTATTAATCTTCAGAGATGTATCTTAATTAAGAAATTAAGT-----TCTTTAAAAA	1833

Db	112	GGACACAGGGAGCAGCCGAGCAGCGGTTTCCCGCAACCCGATACATCGACAGAAATTTCT	171
Qy	121	CGGCCCCAGCCCGCGGGGAGATCTCTGAAAAGATGGCTACAGAACATGTTAATGAAT	180
Db	172	CGGCTCAGCCCAACGGGGAGATCTCTGGAAACATGGCTACAGAACATGTTAATGAAT	231
Qy	181	GGTACTGAAGGCCCATGAGTACTACTCTTCGACGTTATCCATTCAAGAAATTTTCAGACA	240
Db	232	GGTACTGAAGGCCCATGAGTACTACTCTTCGACGTTATCCATTCAAGAAATTTTCAGACA	291
Qy	241	TTGCTGATGCTGGTTTACCAAGAAAGTTGCTGAAAACCTAGTAATTTACGTGCA	300
Db	292	TTGCTGATGCTGGTTTACCAAGAAAGTTGCTGAAAACCTAGTAATTTACGTGCA	351
Qy	301	GGGCTAGTTGCAATAGTATTTAGATGAAGAAGCTATGAGAGCTTTAAAGATTCAT	360
Db	352	GGGCTAGTTGCAATAGTATTTAGATGAAGAAGCTATGAGAGCTTTAAAGATTCAT	411
Qy	361	GAAAGCGGCGCATTTGGCAGTGCTTCAACAGTTTAAAGACATGATCTCTCATGTTGAG	420
Db	412	GAAAGCGGCGCATTTGGCAGTGCTTCAACAGTTTAAAGACATGATCTCTCATGTTGAG	471
Qy	421	AAACAAAGTGCCTTTTATATGGAGTATGAAGACTTACAGGAGAGAAAACAAGGG	480
Db	472	AAACAAAGTGCCTTTTATATGGAGTATGAAGACTTACAGGAGAGAAAACAAGGG	531
Qy	481	ACCAAAGTACAGACTCTAGTAAAGACACAGATAGAGCAAAAGATTAAAGCACTTTGGAA	540
Db	532	ACCAAAGTACAGACTCTAGTAAAGACACAGATAGAGCAAAAGATTAAAGCACTTTGGAA	591
Qy	541	AGAACAGGCTACACACTTATGTGATCTACAGGCTCAGAGGAATAGAGAACCACTTCCA	600
Db	592	AGAACAGGCTACACACTTATGTGATCTACAGGCTCAGAGGAATAGAGAACCACTTCCA	651
Qy	601	GATTCGCTTTATTCAGGTCAGCAGGCTTCGTGTCGACCTGAGATATTTGTTGGGAAAGATC	660
Db	652	GATTCGCTTTATTCAGGTCAGCAGGCTTCGTGTCGACCTGAGATATTTGTTGGGAAAGATC	711
Qy	661	CCACAGAGATCTGTTTGAGAGTATAGCTTGTTCCATTATTTGAGAAAGCTGACCTTATGCG	720
Db	712	CCACAGAGATCTATTTGAGAGTATAGCTTGTTCCATTATTTGAGAAAGCTGACCTTATGCG	771
Qy	721	GATCTTCGTTTATATGATGATCGGCTCAGTGTCTCAACAGAGTTATCGTTGTCACT	780
Db	772	GATCTTCGTTTATATGATGATCGGCTCAGTGTCTCAACAGAGTTATCGTTGTCACT	831
Qy	781	TTTTTGTATAAAAGAACGACACAAAGAGGCTGTTAACTGTATATATCATGTAATTTGCT	840
Db	832	TTTTTGTATAAAAGAACGACCTCAGAGAGGCTGTTAACTGTATATATCATGTAATTTGCT	891
Qy	841	TCCGGGAGCACAATTGGTGTCTGCATCTCAGTTGTCACAAATAGGCTTTTGTGGGCTCG	900
Db	892	TCTGGAAAAACATTTGGTGTCTGCATCTCAGTTGTCACAAATAGGCTTTTGTGGGCTCT	951
Qy	901	ATTCCTTAAGAGTAAACCAAGAGACAGATTTCTTGAGGAATTTAGCAAAAGTGAACAGAGGT	960
Db	952	ATTCCTTAAGAGTAAACCAAGAGAACAGATTTCTTGAGGAATTTAGCAAAAGTGAACAGAGGT	1011
Qy	961	CTCACAGATGTCAATTTTATACACCAACTGATGACAAAGAAAAAAACAGAGGCTTTTGC	1020
Db	1012	CTTACAGCGTCATTTTATACACCAACCGGATGACAAAGAAAAAAACAGAGGCTTTTGC	1071
Qy	1021	TTTCTTGAATATGAAGATTCACAAACAGCTGCCAAGCAAGAGCTGAGCTATAGTGTGT	1080
Db	1072	TTTCTTGAATATGAAGATTCACAAACAGCTGCCAAGCAAGAGCTGAGCTATAGTGTGT	1131
Qy	1081	AAAGTCAAAAGTCTGGGGAAATGTTGGAATCTTGAATGGCTGATCTTATGAAGATCT	1140
Db	1132	AAAGTCAAAAGTCTGGGGAAATGTTGGAATCTTGAATGGCTGATCTTATGAAGATCT	1191
Qy	1141	GATCTGAAGTTATGCGAAAGATTAAGTGTGTTGTATGCAACTTTCGCAACGGTA	1200

Db	1132	GATCTGAGGTTATGSCAAGATTAAGTCTGTTTGAACGACCTTGCCAAATCTGTA	1251
Qy	1201	ACAGAAAGAAATTTTGAAGAAAGTCAATTAGTCAGTTTGGGAAACGTGAAAGATTAAGAAAG	1260
Db	1252	ACAGAAAGAAATTTTGAAGAAAGCAATTTAGTCAGTTTGGGAAACGTGAAAGATTAAGAAAG	1311
Qy	1261	CTAAAGATTTATGCTTTCATTTCAATTTGATGAGAGATGCTGCTCAAGCTATGGA	1320
Db	1312	TTTAAAGATTTATGCTTTCATTTCAATTTGATGAGAGATGCTGCTCAAGCTATGGA	1371
Qy	1321	GAATGAATGTTAAAGCTTGAGAGGAGAAATATTTGAATTTGTTTGTCTAAAGCACA	1380
Db	1372	GAATGAATGCTGAAGATCTTGAGAGGAGAAATATTTGAATTTGTTTGTCTCAAGCACA	1431
Qy	1381	GATCAGAAAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGAAATCAATGTATGAT	1440
Db	1432	GATCAGAAAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAGCAAGAAATCAATGTATGATC	1491
Qy	1441	GATTCTATCTATTTATGCTTCCATCTCATATGCTCTCCCCCAACAGAGCTGAGGGGTGGA	1500
Db	1492	GATTCTATCTATTTATGCTTCCATCTCATATGCTCTCCCCCAACAGAGGCTGAGGGGTGGA	1551
Qy	1501	GGTAGAGGTGCTATGATATCTCCAGATTTATATGATATGAAATTTATATATAT	1560
Db	1552	GGTAGAGGTGCTATGATATCTCCAGATTTATATGATATGAAATTTATATATAT	1611
Qy	1561	TATGTTATGATTTACCATTAATCTATGCTGAGATATGAAATCCATTAATGTTATGAA	1620
Db	1612	TATGTTATGATTTACCATTAATCTATGCTGAGATATGAAATCCATTAATGTTATGAA	1671
Qy	1621	GATTTTCAAGTTTGAAGCTATGAGGAAAGGGGTGTGAGAGCAAGGGGTGCTGCCATCC	1680
Db	1672	GATTTTCAAGTTTGAAGCTATGAGGAAAGGGGTGTGAGAGCAAGGGGTGCTGCCATCC	1731
Qy	1681	AGAGGTGCTGGGGCTGCTCTCCCGCTGTAGAGCCGTTATTTCAACAAGAGAGCCCT	1740
Db	1732	AGAGGTGCTGGGGCTGCTCTCCCGCTGTAGAGCCGTTATTTCAACAAGAGAGCTCT	1791
Qy	1741	GGATCAGCAAGAGCGCTTCCGCTGTCGAGAGAGAGGTGCCAACAACAAGAGCCGCG	1799
Db	1792	GGATCAGCAAGAGCGCTTCCGCTGTCGAGAGAGAGGTGCCAACAACAAGAGCCGCG	1851
Qy	1800	---GGGAAAGGGGCTGAGAGCGGCTCGAAGCTGTTCAATGAAGCTGATCTGATGT	1857
Db	1852	CAGGAAAGGGGCTGAGAGCGGCTCGAAGCTGTTCAATGAAGCTGATCTGATGT	1910
Qy	1858	GGGATTTACACGAGAGCTGTGAGGTGAGTAAATGTTAAAGAAATCAAGCAACTTTAAATA	1917
Db	1911	GGGATTTACACGAGAGCTGTGAGGTGAGTAAATGTTAAAG--AAATACAGCAACTTTAAATA	1969
Qy	1918	TCTCGAGATGATAGAGCAATTTCTATTGTCAGAGAACCTCTCTATGAAGATATGGAATC	1977
Db	1970	TGTCGGCTGTATAGAGCAATTTCTATTGTCAGAGAACCTCTCTATGAAGATATGGAATC	2029
Qy	1978	AAATACGGGACATTTGAACCTAATATCTTGAGACTTGTATGAATTTCTTAAACAATTTCTC	2037
Db	2030	AAATACGGGACATTTGAACCTAATATCTTGAGACTTGTATGAATTTCTTAAACAATTTCTC	2089
Qy	2038	TGCACTGCAAGTTATTAACCTAAAGCTACTCTATTTTCCAATGTGTTCCA--AAAAATC	2096
Db	2090	TGCACTGCAAGTTATTAACCTAAAGCTACTCTATTTTCCAATGTGTTCCAAGAAATC	2149
Qy	2097	CTTCATTACTTCTAGCATGCTATCTTATTAAGATTAAGTTGTTCTCTTTAAAAA	2153
Db	2150	CTTCATTACTTCTAGCATGCTATCTTATTAAGATTAAGTTGTTTAAAAATCAAA	2206

RESULT 5	BC004001	LOCUS	DEFINITION	ACCESSION	VERSION
	BC004001	2195 bp	mRNA	linear	ROD 07-AUG-2002
			Mus musculus, clone IMAGE:3345242, mRNA, partial cds.		
	BC004001.1	GI:13278369			

QY 2846 GGATATTTTGAAGTTTTCATTAATCATAGAAATTAAGTATATAGCTGTACTTACCTT 2905
DB 1380 GGATATTTTGAAGTTTTCATTAATCATAGAAATTAAGTATATAGCTGTACTTACCTT 1439
QY 2906 TTAGTAAAGATAGTTTCTTATAGTATGTTTATATGCTATCATATAAATGTA 2965
DB 1440 TTAGTAAAGATAGTTTCTTATAGTATGTTTATATGCTATCATATAAATGTA 1439
QY 2966 ATCAGCTTTCATTTCTTATAGTATGTTTATATGCTATCATATAAATGTA 3025
DB 1500 ATCAGCTTTCATTTCTTATAGTATGTTTATATGCTATCATATAAATGTA 1559
QY 3026 TGTATGTTTATTTTCTTAAAGACCTTATGATGATAGTATGATGATGATGAT 3085
DB 1560 TGTATGTTTATTTTCTTAAAGACCTTATGATGATAGTATGATGATGATGATGAT 1619
QY 3086 CCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3145
DB 1620 CCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
QY 3146 CAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3205
DB 1680 CAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1739
QY 3206 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3265
DB 1740 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
QY 3266 ACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3325
DB 1800 ACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1859
QY 3326 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3385
DB 1860 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919
QY 3386 AATGTTTATATAC 3398
DB 1920 AATGTTTATATAC 1932

RESULT 6
LOCUS AR177851 2079 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 2 from patent US 6313266.
ACCESSION AR177851
VERSION AR177851.1 GI:17920206
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Bandman, O., Yue, H., Corley, N. C. and Shah, P.
TITLE Human nucleolin-like protein
JOURNAL Patent: US 6313266-A 2 06-NOV-2001;
FEATURES
source 1..2079
/organism="Unknown"
BASE COUNT 639 a 377 c 535 g 528 t
ORIGIN

Query Match 52.9%; Score 1826.2; DB 6; Length 2079;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 1951; Conservative 0; Mismatches 78; Indels 38; Gaps 4;

QY 2 CGGCTTTGCTCCAGCGGCTGAGCTTGGCGGCAATTTTCAACAGCTCCAGCGCC 61
DB 50 CGGCTTTGCTCCAGCGGCTGAGCTTGGCGGCAATTTTCAACAGCTCCAGCGCC 109
QY 62 GGAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 110 GGAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 169

QY 121 CGGCCCCAGCCCCCGGGGAGATCTCTGGAAGATGGCTTACAGAACTGTAAAT 180
DB 170 CGGCTCAGGCCCAAGGGGAGATCTCTGGAAGATGGCTTACAGAACTGTAAAT 229
QY 181 GGTACTGAAGAGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 230 GGTACTGAAGAGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289
QY 241 TTGCTGATGCTGCTTACCAAGAAAGTTGCTGAAGAACTGATGAAATTTAGTTGCA 300
DB 290 TTGCTGATGCTGCTTACCAAGAAAGTTGCTGAAGAACTGATGAAATTTAGTTGCA 349
QY 301 GGGCTGTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 350 GGGCTGTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
QY 361 GAAGACGGGCAATTTGGCAGTCTTCAACAGTTTAAAGACAGTATCTCTCATGTTGAG 420
DB 410 GAAGACGGGCAATTTGGCAGTCTTCAACAGTTTAAAGACAGTATCTCTCATGTTGAG 469
QY 421 AACAAAGTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 470 AACAAAGTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 529
QY 481 ACCAAAGTACACATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 530 ACCAAAGTACACATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 589
QY 541 AGAACAAGGCTACACATCTTATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 590 AGAACAAGGCTACACATCTTATGATGATGATGATGATGATGATGATGATGATGAT 649
QY 601 GATTCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 650 GATTCGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 709
QY 661 CCGAGATCTGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 710 CCGAGATCTGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
QY 721 GATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 770 GATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
QY 781 TTTTGTACAAAG 840
DB 830 TTTTGTACAAAG 889
QY 841 TCCGGGAG 900
DB 890 TCCGGGAG 949
QY 901 ATTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 950 ATTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
QY 961 CTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1010 CTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
QY 1021 TTTTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1070 TTTTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
QY 1081 AAAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1130 AAAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
QY 1141 GATCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1190 GATCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249

OY	1201	ACAGAAAGAAATTTTAAAGAAAGTCAATTAAGTCACTTTGGAAACATCGAAGCAAGAAAGAAAG	1260
Db	1250	ACGAGAGAGATTTTAAAGAAAGCAATTTAGTCACTTTGGAAACATCGAAGCAAGAAAGAAAG	1309
OY	1261	CTAAAAGATTAATGCTTTCATTTCAATTTGATGAGAGAGATGCTGCTCAAGCTATGAGAA	1320
Db	1310	TTAAAAAGATTAATGCTTTCATTTCAATTTGATGAGAGAGATGCTGCTCAAGCTATGAGAA	1369
OY	1321	GAATGAATATGTTAAACATTTGGAGGAGAAATATTTGAAATTTGTTTTGCTAAAGCAACA	1380
Db	1370	GAATGAATATGCTAAACATTTGGAGGAGAAATATTTGAAATTTGTTTTGCTAAAGCAACA	1429
OY	1381	GATCAAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGCAAGCAATCAATGTTATGAT	1440
Db	1430	GATCAAGAAAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGCAAGCAATCAATGTTATGATC	1489
OY	1441	GATTAATCAATTAATATGCTCACTCATATGCTCTCCCAACAAAGAGTGAAGGCGGTGA	1500
Db	1490	GATTAATCAATTAATATGCTCACTCATATGCTCTCCCAACAAAGAGTGAAGGCGGTGA	1549
OY	1501	GCTAAGAGTGGCTATGATGATTCCTCAGATTAATTAATGATTAAGAAATTAATTAATGAT	1560
Db	1550	GCTAAGAGTGGCTATGATGATTCCTCAGATTAATTAATGATTAAGAAATTAATTAATGAT	1609
OY	1561	TATGTTATGATTAACCAATACTATCGTGGTGAATATGAAGATCCATATGTTATGAA	1620
Db	1610	TATGTTATGATTAACCAATACTATCGTGGTGAATATGAAGATCCATATGTTATGAA	1669
OY	1621	GATTTTCAAGTTGGAGCTAGAGAGAAAGGAGTGGTAAGAGAGCAAGGGGTCTGCTCATCC	1680
Db	1670	GATTTTCAAGTTGGAGCTAGAGAGAAAGGAGTGGTAAGAGAGCAAGGGGTCTGCTCATCC	1729
OY	1681	AGAGGTCGTGGGAGCTGCTCTCCCGCTGGTAAAGACCGGTTATTCACAGAGAGAGCGCT	1740
Db	1730	AGAGGTCGTGGGAGCTGCTCTCCCGCGGTAAAGACCGGTTATTCACAGAGAGAGGTCCT	1789
OY	1741	GGATCAGCAAGAGCGCTTGCAGGTGCGAAGAGGTGCCAACACAAAGAGCGCGGAG	1800
Db	1790	GGATCAGCAAGAGCGCTTGCAG-----CAG 1814	
OY	1801	GGAAAAAGGGGTGAGAGCGCGGTCTGACCTGTTACATTAAGAGCTGACTTATGTTGGG	1860
Db	1815	GGAAAAAGGGGTGAGAGCGCGGTCTGACCTGTTACATTAAGAGCTGACTTATGTTGGG	1873
OY	1861	ATTAAACCAAGAGCTTGCAGTGAAGTAAATGTAAGAAAAATCAAGCAACTTAATATCT	1920
Db	1874	ATTAAACCAAGAGCTTGCAGTGAAGTAAATGTAAG- AAATCAAGCAACTTAATATATCT	1932
OY	1921	CGGAGTATTAAGAGCAATTTCAATTGCAAGAAACCTCTTAAGAAATCATGGAATCAAA	1980
Db	1933	CGGCTGTATTAAGAGCAATTTCAATTGCAAGAAACCTCTCTAAGAAATCATGGAATCAAA	1992
OY	1981	TACGGGACATTTGAATTAATCTTGAACCTTTGTTATGAATTTCTTTAAACAATTTCTCTGC	2040
Db	1993	TACGGGACATTTGAATTAATCTTGAACCTTTGTTATGAATTTCTTTAAACAATTTCTCTGC	2052
OY	2041	AGTGAAGTTATTAACCTAAAGCTACT 2067	
Db	2053	AGTGAAGTTATTAACCTAAAGCTACT 2079	

RESULT 7	
HSJ775C13	
LOCUS	
DEFINITION	HSJ775C13 103819 bp DNA linear PRI 10-MAR-2001
	Human DNA sequence from clone RP4-775C13 on chromosome 20p12.1-13.
	Contains an hnRNP R (heterogeneous nuclear riboprotein R, RRM ma
	binding protein GRY-RBP) pseudogene, ESTs, STS and GSS, complete
	sequence.
ACCESSION	AL109618
VERSION	AL109618.1 GI:5668665
KEYWORDS	HTG; GRY-RBP; hnRNP; RRM.
SOURCE	human.
ORGANISM	Homo sapiens

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE         1 (bases 1 to 103819)
JOURNAL      Matthews, L.
              Direct Submission
              Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA. E-mail enquiries: humquerry@sanger.ac.uk clone
              requests: clonerequests@sanger.ac.uk
              During sequence assembly data is compared from overlapping clones.
              Where differences are found these are annotated as variations
              together with a note of the overlapping clone name. Note that the
              variation annotation may not be found in the sequence submission
              corresponding to the overlapping clone, as we submit sequences with
              only a small overlap as described above.
              This sequence was finished as follows unless otherwise noted: all
              regions were either double-stranded or sequenced with an alternate
              chemistry or covered by high quality data (i.e., phred quality >=
              30); an attempt was made to resolve all sequencing problems, such
              as compressions and repeats; all regions were covered by at least
              one plasmid subclone or more than one M13 subclone; and the
              assembly was confirmed by reassociation primary accession numbers given
              in the feature table with their source databases: Emr, EMBL, Sw;
              SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
              database can be found at
              http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
              was generated from part of bacterial clone contigs of human
              chromosome 20, constructed by the Sanger Centre Chromosome 20
              Mapping Group. Further information can be found at
              http://www.sanger.ac.uk/HGP/Chr20
              RP4-775C13 is from the library RPCT-4 constructed by the group of
              Pieter de Jong. For further details see
              http://www.chori.org/bacpac/home.htm
              VECTOR: pCYPAC2
              This sequence is the entire insert of clone RP4-775C13.

FEATURES
  source
    location/Qualifiers
      1..103819
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="20"
        /map="p12.1-13"
        /clone="RP4-775C13"
        /clone_1lb="RPCT-4"
      1..236
        /note="ANY repeat: matches 1..236 of consensus"
      293..562
        /note="LM4C repeat: matches 1844..1624 of consensus"
      819..1471
        /note="LM3C repeat: matches 969..1634 of consensus"
      2177..2588
        /note="LM3C repeat: matches 14..431 of consensus"
      2593..2652
        /note="LM3C repeat: matches 474..533 of consensus"
      2701..3176
        /note="LM3C repeat: matches 474..533 of consensus"
      2814..2978
        /note="15 copies 2 mer tt 54% conserved"
      2996..3196
        /note="15 copies 11 mer 66% conserved"
      3031..3184
        /note="67 copies 3 mer tcc 56% conserved"
      3296..4082
        /note="14 copies 11 mer 63% conserved"
      4197..4315
        /note="LM4 repeat: matches 3350..4128 of consensus"
      4316..4738
        /note="LM4 repeat: matches 3021..3136 of consensus"
      4739..5118
        /note="MST1 repeat: matches 1..426 of consensus"
      5146..5288
        /note="LM4 repeat: matches 2626..3021 of consensus"
      5345..6619
        /note="match: GSS: Emr:B90786"
      5345..6619
        /note="LM4C repeat: matches 3863..5135 of consensus"

```


OY	362	AAGAGGGGCAATGGCAGTGGTCCAAACAGTTTAAAGACAGTACCTCTCTCAAGTTTCA	421
Db	7559	AAGAGGTGCAATGGCAGTCTTTCACAG--TTAAAGCAGTATCTCTCGTGTCCAG	7617
OY	422	ACAAAGTGCCTTTTATATGTGAGTCATGAAAGCTTACAGGCAAGAGAAACAGGGGA	481
Db	7618	ACAAAGTGCCTTTTATATGTGAGTCATGAAACCTTATAGCAGAGAAACAGAGGA	7677
OY	482	CCAAAGTACGACTTAATGTAAGAGCAGATGAGGCAAAAGTTAAGGCATTTTGGAAA	541
Db	7678	CCAAAGTACGACTTATGTAAGAGCAGATGAGGCAAAATTTAAGGCATCTTGGAAA	7737
OY	542	GAAAGGCTACACCTTGAATGTGCTACAGGTGAGGAAAGTATGAGAGCACCTCCAG	601
Db	7738	GAGCAGCTACACCTTGAATGTGCTACAGGACAGGAAATATGAGAGCACCTCCAG	7797
OY	602	ATTCGTTTATTCAGTCCAGGCTCTCTTGGACCTAGATATTTGGGGGAAGATCC	661
Db	7798	ATTCGTTTATTCAGGCTCAGAGGCTCTCTTGGACCTAGATATTTGGGGGAAGATCC	7857
OY	662	CGAGATCTGTGTTGAGGATGAGCTGTTCATATTTAGAAAGCTGACCTATATGGG	721
Db	7858	CAAGATCTATTTAGAGATGAACCTGTTCATATTTAGAAAGCTGACCTATATGGG	7917
OY	722	ATCTTCGTTTAAATGATGATCCGCTCAGCTGCTCACAAGGTTATGCTTGTGCATT	781
Db	7918	ATCTTCGCTTAAATGATGATCCCTCAGCTGCTCACAATGAGTTATGCTTGTGCATT	7977
OY	782	TTTGTACAAAGAGCAGCACAAGAGGCTGTAACTGTATATATATCATGAATTGGT	841
Db	7978	TTTGTACAAAGAGCAGCTCCGAGGCTG--TTAAACTGTATATATCATGAATTGGT	8035
OY	842	CCGGAGCACAATTGGTGTCTGCATCTCAGTTGCCAACATATAGCTTTTGTGGCTCGA	901
Db	8036	CTGAAACACATGGTGTCTGCATCTCAGTTGCCAACATATAGGCTTTTGTGGCTCGA	8095
OY	902	TTTCTTAAGAG--TAAACCAAGAGCAGATCTTGAGAAATTATGAC--AAAGTACAGAGG	959
Db	8096	TTTCTTAAGAGTTAAACCAAGAGCAGATCTTGAGAAATTATGACAAAGTACAGAGG	8155
OY	960	TCTCAGAGATGCATTTTATATACCAACCACTGTATGACAAGAAAAAACAAGGCTTTTG	1019
Db	8156	TCTTACAGACTCATTTTATATACCAACCGGATGACAGAAAAAACAAGGCTTTTG	8213
OY	1020	CTTTCTTGAATATGAGATGACAAAACAGCTGCCAGGACAACGTAGCTATATGAGTG	1079
Db	8214	CTTTCTTGAATATGAGATGACAAAACAGCTGCCAGGACAACGCTATATGAGTG	8273
OY	1080	TAAAGTCAAAGTCTGGGGAATATGTGAACTGTGATGAGGCTGATCTATTAAGATCC	1139
Db	8274	TAAAGTCAAAGTCTGGGGAATATGTGAACTGTGATGAGGCTGATCTATTAAGATCC	8333
OY	1140	TGATCTCGAAGTTATGCGCAAGGTAAAGTCGTTTATACCAACCTTGGCCACAACGCT	1139
Db	8334	TGATCTCGAAGTTATGCGCAAGGTAAAGTCGTTTATACCAACCTTGGCCACAACGCT	8393
OY	1200	AACAGAGAAATTTTACAAAAGTCAATTAGTCAGTTTGGGAAAACGTGAATAAGAA	1259
Db	8394	AACGAGAGAGTTTATAGAAAAGCATTTAGTCAAGTTTGGGAAAACGTGAATAGTGA	8453
OY	1260	GCTAAAGATTAATGCTTTCATTCATTTTGAATGAGAGATGCTGCTCAAGGCTATGGA	1319
Db	8454	GTTAAAGATTAATTCATTCATTCATTTTGGTGAATGAGATGCTGCTCAAGGCTATGGA	8513
OY	1320	AGAAATGATATGTTAAGACTTTGAGGGGAAATATTTGAATTTGTTTGTCTAAGCCACC	1379
Db	8514	AGAAATGATATGTTAAGACTTTGAGGGGAAATATTTGAATTTGTTTGTCTAAGCCACC	8573
OY	1360	AGATCAGAGAGGAAAGAAAGAAAGCTCAGAGGCAAGACGAAAGATCAATTTATGAA	1439
Db	8574	AAATCAGAAAGGAAAGAA--AAAGCTCAGAGGCAAGACGAAAAATCAAAATTTATGA	8631
OY	1440	TGATTAATCAATTAATGCTCACCTCATATGCTCCCCCAACAGATGCGAGGCGTGG	1499

[illegible]

```

RESULT 8
AL390737
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL390737 212659 bp DNA linear HTG 21-MAR-2001
Homo sapiens chromosome 13 clone RP11-190P17, *** SEQUENCING IN
PROGRESS ***, 16 unordered pieces.
AL390737
AL390737.9 GI:13398796
HTG; HTGS _PHASE1; HTGS _DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 212659)
Burton,J.
Direct Submission
Submitted (17-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Mar 21, 2001 this sequence version replaced gi:13396612.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
-----

```

Query Match	Best Local Similarity	Matches 2014, Conservative	52.2%	Score 1802;	DB 2;	Length 212659;
QY 3	GGGTTTGTCTGACGCGGCTGAGCTTGCGGCGCATTTTAAACAAGCTCCACTGCGCGG	62				
DB 82955	GGGTTTGTCTGACGCGGCTGAGCTTGCGGCGCATTTTAAACAAGCTCCACTGCGCGG	83014				
QY 63	GACACAGGAGCAGCGACACGCGTCCCGCCAGCCGCGCGAG-CGATTCCTC	121				
DB 83015	GACACAGGAGCAGCGACACGCGTCCCGCCAGCCGCGCGAG-CGATTCCTC	83074				
QY 122	GGCCCCAGCCCCGGGGGAGATCTCTGGAAAACATGCTACAGAACTGTTATGGAATG	181				
DB 83075	CCCCCATCTCCACGGGGGAGATCTTGTGGAAAACATGCTACAGAACTGTTATGGAATG	83134				
QY 182	GTACTGAGAGCGCATGTGATCTACTCTTCTGAGAGTTATCCATTCAGAAAATTTTCAACAT	241				
DB 83135	ATACTGAGAGCGCATGTGATCTACTCTTCTGAGAGTTATCCATTCAGAAAATTTTCAACAT	83194				
QY 242	TGCTTGATGCTGTTTACACAGAAAGTGTCTGAAAACCTAGATGAATTTACGTTGACG	301				
DB 83155	TGCTTGATGCTGTTTACACAGAAAGTGTCTGAAAACCTAGATGAATTTACGTTGACG	83254				
QY 302	GGCTAGTTGCACTATGATGATTTAATGAAGAAGCTATCGAAGCTTTAAAGAATTCATG	361				
DB 83255	GGCTAGTTGCACTATGATGATTTAATGAAGAAGCTATTCAGAACTTTAAAGAATTCATG	83314				
QY 362	AAGAGCGGCGCATTTGCGCATGCTTCAACAGTTTAAAGACATGATCTCTCATATGTTACA	421				
DB 83315	AAGAGCGGCGCATTTGCGCATGCTTCAACAG-TTAAAGACATGATCTCTCATATGTTACA	83373				
QY 422	ACAAAAGTCCCTTTTATGTGAGATCACTGAAGCTTACAGGCGAGAGAAAACAGGGGA	481				
DB 83374	ACAAAAGTCCCTTTTATGTGAGATCACTGAAGCTTATAGGCGAGAGAAAACAGGGGA	83433				
QY 482	CCAAAAGTACAGACTCTATGTAAGAACCCAGATGAGGACCAAGATTTAAGGCACTTTGGAA	541				
DB 83434	CCAAAAGTACAGACTCTATGTAAGAACCCAGATGAGGACCAAAATTTAAGGCACTTTGGAA	83493				
QY 542	GAAACGGCTACACACTTATATGTGACTACAGGTCTAGAGAAATATAGAGAACCACTCCAG	601				
DB 83494	GAAACGGCTACACACTTATATGTGACTACAGGTCTAGAGAAATATAGAGAACCACTCCAG	83553				
QY 602	ATTCGTTTATTCAGGTGAGAGGCTTCTGTGGCACTGAGATATTTGTGGGGAAGATCC	661				
DB 83554	ATTCGTTTATTCAGGTGAGAGGCTTCTGTGGCACTGAGATATTTGTGGGGAAGATCC	83613				
QY 662	CCAGAGATCTGTTGAGGATGAGCTTGTTCCATTAATTTGAAGAAAGCTGACCTTAATGG	721				

Db	84687	TGGATCAGCAAGAAGCCGTTTGAAGTGCCAGAGAGAGGTGCCAACAACAAAGAAGCGCGG	84746
Oy	1800	---GGGAAAAGGGGTCGAGGCGCGTCTGACCTGTTAACAATGAAGACTGTGCATTG	1856
Db	84747	GCAAGGAAAAAGGGGTCGAGGCTGGTCTTGACCTGTTAACAATGAAGACTGTGGA-TG	84805
Oy	1857	TGGGATTACACCAGAAGCTTGCAGTGGAGTAATGATAAGAAAATCAAGCACTTTAAT	1916
Db	84806	TGGGATTACACCAGAAGCTTGCAGTGGAGTAATGATAAG- AAATCAAGCAACTTTAAT	84864
Oy	1917	AATCGGANTGTATAGACCATTTCTAATTGCAAGAACCCCTCTATGAAGATCARGAAT	1976
Db	84865	AATCGGCTGTATAGAGCATTTCTATTGCAAGAACCCCTCTGTGAAGATCARGAAT	84924
Oy	1977	CAATACCGGACATTTGAAGTAACTAATCTGTGGAAGCTTTGTATGAATTTCTTTAACAATTTCT	2036
Db	84925	CAATACCGGACATTTAATCTAATCTGTGGAAGCTTTGTATGAATTTCTTTAACAATTTCT	84984
Oy	2037	CTGCAGTGCAGATTATTAATACTAAGACTATCTATTTCCAATGTGTCCAA-AAAAAT	2095
Db	84985	CTGCAGTGCAGATTATTAATACTAAGACTATCTATTTCCAATGTGTCCAAGAAT	85044
Oy	2096	CCTTCATACCTTCAGCANGTAGATCTTTAATAAGAAATAAGTGTCTCTTTAAA	2151
Db	85045	CCTTCATACCTTCAGCANGTAGATCTTTAATAAGAAATAAGTGTCTTTAAAAAAAA	85100
RESULT 9			
AF037448			
LOCUS	AP037448	2932 bp	mRNA linear PRI 08-APR-1996
DEFINITION	Homo sapiens RRM RNA binding protein Gry-rbp (GRY-RBP) mRNA,		
ACCESSION	AF037448		
VERSION	AF037448.1	GI:3037012	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2932) Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B. Cloning of a RRM RNA binding protein wltch glycine, arginine and tyrosine-rich C-terminus Unpublished		
JOURNAL	2 (bases 1 to 2932)		
REFERENCE	Du G., Yuan J., Pan M., Yao H., Chen J. and Qiang B.		
AUTHORS	Direct Submission		
TITLE	Submitted (09-DEC-1997) Department of Biochemistry and Molecular Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China		
JOURNAL	location/Qualifiers		
AUTHORS	1 . 2932		
TITLE	/organism="Homo sapiens"		
JOURNAL	/db_xref="taxon:9606"		
AUTHORS	1 . 2932		
TITLE	/gene="GRY-RBP"		
JOURNAL	526 . 2397		
AUTHORS	/gene="GRY-RBP"		
TITLE	/note="RRM RNA binding protein; RNA recognition motif binding protein; Glycine, arginine and tyrosine-rich at C-terminus"		
JOURNAL	/codon_start=1		
AUTHORS	/product="Gry-rbp"		
TITLE	/protein_id="AACI2926.1"		
JOURNAL	/db_xref="GI:3037013"		
AUTHORS	/translation="MATEHVNGSTPEPDTTSVHSNFPQTLLDGLPKVAEKLE EIVAVIGASVSDLDIERALEAKENEDGALAVLOFQSDSLSHVNKSAFLSGVMKVITY RORKGVASTDASSKGDEAKIKALERTGYLDDVTVGORVGPSPDSVGOQSV GTELVFKIPDLDELDELVEPEKAGIPMDLRMDPLDTGLNRGAYATFCKEAADG AVTLVNNHEIRSGDEIGVCI SVANNRLPVGISPSKTKBOILFRFSVTEGLTVYLIT HQPDKKKNSFCFLIEYEDHKTAQAARRLSMGKYKWGNVGTETWADPIEDPPETWA AKKVLVTRNLANTVIELKEKASFGKLERVKVDVAFIHDERDGAVKAMEENNNA GLKEGINIIIVFAKPDPDKRKERKAQRQAOKAMQMDYYIYGPPNPMPPRGGRGGG		

RGYGYPDYVYEDYDYDYHNYRGYEDPYGYEDFOYGARGRGARGAP
SRGGAAPRGARGYSGORGSARGAGAGAGAGAGAGAGAGAGAGAGAG
ADGYNOBDSRRORNTNNQNMGSORIPADPLOGGHSNRYGKSNORFYODTRQOMK*

BASE COUNT 868 a 520 c 823 g 721 t

ORIGIN

Query Match 48.5%; Score 1674; DB 9; Length 2932;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 85; Indels 1; Gaps 1;

QY 2 CGGGTTTGTGCGACGGCGGTGAGCTTGCGCCGCAATTTTACAACAGCTCCACGCGCC 61
Db 373 CGGGTTTGTGCGACGGCGGTGAGCTTGCGCCGCAATTTTACAACAGCTCCACGCGCC 432
QY 62 GGACACAGGGGACGACGACGCGTCCCGGACGCGACCGCGGTGAGAC-GGATTCCT 120
Db 433 GGACACAGGGGACGACGACGCGTCCCGGACGCGACCGGACGCGGTGAGACGAGATTCCT 492
QY 121 CGGCCCCAGCCCGCGGGAGATCTCTGAAAACATGGCTACAGAACATGTTAATGAAAT 180
Db 493 CGGCTCAGCCCAACGGGGAGATCTCTGAAAACATGGCTACAGAACATGTTAATGAAAT 552
QY 181 GGTACTGAAGAGCCCATGATGATCTTCAAGAGTTATCCATTCAGAAAATTTTCAGACA 240
Db 553 GGTACTGAAGAGCCCATGATGATCTTCAAGAGTTATCCATTCAGAAAATTTTCAGACA 612
QY 241 TTGCTTATGCTGGTTTACCAAGAAAGTTGCTGAAAACATGATGAAATTTTACGTTGCA 300
Db 613 TTGCTTATGCTGGTTTACCAAGAAAGTTGCTGAAAACATGATGAAATTTTACGTTGCA 672
QY 301 GGGCTAATGTCACATAGTATGTTTATGATGAAAAGCTATCGAAGCTTTTAAAAGATTCAT 360
Db 673 GGGCTAATGTCACATAGTATGTTTATGATGAAAAGCTATCGAAGCTTTTAAAAGATTCAT 732
QY 361 GAAGACGGCGCATGTCGACGCTTCAACAGTTTAAAGACAGTATCTCTCATGTTTCAAG 420
Db 733 GAAGACGGCGCATGTCGACGCTTCAACAGTTTAAAGACAGTATCTCTCATGTTTCAAG 792
QY 421 AACAAAAGTGCCTTTTATGATGAGTCAATGAAAGCTTACAGGCGAGAGAAAACAGGGG 480
Db 793 AACAAAAGTGCCTTTTATGATGAGTCAATGAAAGCTTACAGGCGAGAGAAAACAGGGG 852
QY 481 ACCAAAGTAGACAGCTAGTAAAGACAGATGAGCAAAAGATTAAAGCACTTTTGGAA 540
Db 853 ACCAAAGTAGACAGCTAGTAAAGACAGATGAGCAAAAGATTAAAGCACTTTTGGAA 912
QY 541 AGAAGACGGCTACACATTTGATGATCTACAGTCAAGAGAAATGAGAGACCACTTCCA 600
Db 913 AGAAGACGGCTACACATTTGATGATCTACAGTCAAGAGAAATGAGAGACCACTTCCA 972
QY 601 GATTCCGTTTATTCAGGTCAGGAGCGCTTCTGTTGGCACTGAGATTTTGTGGGAATATC 660
Db 973 GATTCCGTTTATTCAGGTCAGGAGCGCTTCTGTTGGCACTGAGATTTTGTGGGAATATC 1032
QY 661 CCGAGAGATCTGTTGAGATGAGCTTGTTCATTATTTTGAAGAAAGCTGACCTAATATG 720
Db 1033 CCGAGAGATCTATTTGAGATGAGCTTGTTCATTATTTTGAAGAAAGCTGACCTAATATG 1092
QY 721 GATCTTCTGTTAATGATGATCCGCTCACTGCTCAACAGAGGTTATGCGTTTGTCACT 780
Db 1093 GATCTTCTGTTAATGATGATCCGCTCACTGCTCAACAGAGGTTATGCGTTTGTCACT 1152
QY 781 TTTTGTACAAAAGAGAGACAGAGAGCTGTTTAACTGTATATATCATGAAATTCGT 840
Db 1153 TTTTGTACAAAAGAGAGAGCTGAGAGAGCTGTTTAACTGTATATATCATGAAATTCGT 1212
QY 841 TCCGGAAGACATTTGCTGCTGATCTCAGTTTSCAAACAATAGAGCTTTTGTGGGCTCG 900
Db 1213 TCCGGAAGACATTTGCTGCTGATCTCAGTTTSCAAACAATAGAGCTTTTGTGGGCTCT 1272
QY 901 ATTCTTAAGAGTAAACCAAGAGACAGATTTCTTGAGAAATTTTACAAAGTGAAGAGGT 960
Db 1273 ATTCTTAAGAGTAAACCAAGAGACAGATTTCTTGAGAAATTTTACAAAGTGAAGAGGT 1332

QY 961 CTCACAGATGCTATTTTATACCAACCACTGATGACAAAGAAAACAGAGCTTTTGC 1020
Db 1333 CTTCACAGATGCTATTTTATACCAACCACTGATGACAAAGAAAACAGAGCTTTTGC 1392
QY 1021 TTTCTTGAATATGAGATGACAAAACAGCTGCGCAGGACGATAGCTAATGATGAT 1080
Db 1393 TTTCTTGAATATGAGATGACAAAACAGCTGCGCAGGACGATAGCTAATGATGAT 1452
QY 1081 AAGTCAGAGTCTGGGAAATGTTGAACTGTTGAGTGGGCTGATCTTATGAAATCCT 1140
Db 1453 AAGTCAGAGTCTGGGAAATGTTGAACTGTTGAGTGGGCTGATCTTATGAAATCCT 1512
QY 1141 GATCTGAGATGATGAGCAAAAGGTTAAAGTCTGTTTATGAGAACTTGGCAACGCTA 1200
Db 1513 GATCTGAGATGATGAGCAAAAGGTTAAAGTCTGTTTATGAGAACTTGGCAACGCTA 1572
QY 1201 ACAGAAAGATTTTATGAAAAGTCAATTAATGCTAGTTTGGAAACTGGAACGATTAAGAAG 1260
Db 1573 ACAGAAAGATTTTATGAAAAGTCAATTAATGCTAGTTTGGAAACTGGAACGATTAAGAAG 1632
QY 1261 CTAAAGATTTATGCTTTCATTCATTTTATGAGAGATGCTGCTCAAGGCTATGAA 1320
Db 1633 CTAAAGATTTATGCTTTCATTCATTTTATGAGAGATGCTGCTCAAGGCTATGAA 1692
QY 1321 GAAATGAATGTTAAAGACTTGGAGAGGAAATATTTGAATTTGTTTGTCTAAGCCACA 1380
Db 1693 GAAATGAATGTTAAAGACTTGGAGAGGAAATATTTGAATTTGTTTGTCTAAGCCACA 1752
QY 1381 GATCAGAAAGGAAAGAAAGAAAGTCTCAGAGCAAGCAGCAAGATCAATGATGAT 1440
Db 1753 GATCAGAAAGGAAAGAAAGAAAGTCTCAGAGCAAGCAGCAAGATCAATGATGAT 1812
QY 1441 GATTACTATTTATGCTCACCCTCATATGCTCTCCCAACAGAGGTCGAGGCGTGA 1500
Db 1813 GATTACTATTTATGCTCACCCTCATATGCTCTCCCAACAGAGGTCGAGGCGTGA 1872
QY 1501 GGTAGAGGTGCTATGAGATATCTCCAGATTTATGATGATAGAGATTTATGATGAT 1560
Db 1873 GGTAGAGGTGCTATGAGATATCTCCAGATTTATGATGATAGAGATTTATGATGAT 1932
QY 1561 TATGTTATGATTAACCATATCATGCTGCTGATGATGAAAGATTCATGATGATGAA 1620
Db 1933 TATGTTATGATTAACCATATCATGCTGCTGATGATGAAAGATTCATGATGATGAA 1992
QY 1621 GATTTCAGAGTTGAGCTAGAGAGAGGCTGTGAGAGAGAGAGGCTGCTCATCC 1680
Db 1993 GATTTCAGAGTTGAGCTAGAGAGAGGCTGTGAGAGAGAGAGGCTGCTCATCC 2052
QY 1681 AGAGGTGCTGAGGCTGCTCTCCCGTGTGAGAGCGGTTATTCACAGAGAGAGGCT 1740
Db 2053 AGAGGTGCTGAGGCTGCTCTCCCGTGTGAGAGCGGTTATTCACAGAGAGAGGCT 2112
QY 1741 GATCAGCAAGAGAGGCTTGGCGGTGAGAGAGAGGCTCCCAACAAGAGAGCGCGGG 1800
Db 2113 GATCAGCAAGAGAGGCTTGGAGGTGAGAGAGAGGCTCCCAACAAGAGAGCGCGGG 2172
QY 1801 GGAAGAGGAGGCTGAGGCGGCTC 1822
Db 2173 GTRACGTGCTGAGAGGAGGCTG 2194

RESULT 10
AF093821
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF093821 2373 bp mRNA linear ROD 04-OCT-1998
Mus musculus RRM RNA binding protein GRV-RBP mRNA, complete cds.
AF093821.1 GI:3694985
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2373)
 AUTHORS Du, G., Zhou, Y., Chen, J., Yuan, J., and Qiang, B.
 TITLE Direct Submission
 JOURNAL Submitted (23-SEP-1998) Department of Biochemistry and Molecular
 Biology, Institute of Basic Medical Sciences, 5 Dong Dan San Tiao,
 Beijing 100005, P. R. China

FEATURES
 Location/Qualifiers

1..2373
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 7..1884
 /codon_start=1
 /product="RBM RNA binding protein GRV-RBP"
 /protein_id="AAC62511.1"
 /db_xref="GI:3694986"
 /translation="MATEHYNGNGTEBEPDITTSANVHSENPOTLLDAGLPQVAKELD
 EIVAGLVASDLDERALIKRPNDGALAVLOPKDSLHVNKSAFLGVMKTY
 RQREKGTVAADSKQPDPAKIKALERTGYLDVYTGKRGPPDSVSGQPSV
 GEIEIVKIPDLFEDLVPLFEKAGPNDLMDPLGLNKGYPFVETKEA
 AOEAAKLNNHIEIRSGKIHGVCISVANNELFVGSIPKSKTQOILEEFKTEGLTDV
 ILVHODDDKKRRGPGFOEYEDHKAQAARRLSKGVNNGVTEWADIEDDP
 EYMAKYKULFVFNILANTVTEHILEKSPSGCLERVKIKDYAFHFERDAVYAME
 ENMGKDLBSENTEIVFAKPPDQKREKKAQKQAAKQMDYIYTGPPHPPTEKRG
 RGKRGYGPEDYGYEDYDYGYDYNHRYGVEDPYGYEDPQVANGRGGRGARG
 AAPSRGGAAPPRGAGYSGRGA PGSGARVGRGAQQRGGRGARGRGANVGK
 RADGYNQDPTKRQTNHQMGSQPIAQQLPGSDHSNRYKSENGEYDPTFGQW
 K"

BASE COUNT 740 a 412 c 587 g 634 t
 ORIGIN

Query Match 46.6%; Score 1608; DB 10; Length 2373;
 Best Local Similarity 98.1%; Pred. No. 0;
 Matches 1652; Conservative 0; Mismatches 20; Indels 12; Gaps 2;

QY 148 GGAACATGCTACAGAACATGTTAATGAAATGTACTGAAGAGCCCATGATACTACT 207
 DB 1 GGAACATGCTACAGAACATGTTAATGAAATGTACTGAAGAGCCCATGATACTACT 60
 QY 208 TCAGAGTTATCATTCAGAAATTTTTCAGACATGCTGATCTGGTTTACCAAGAAA 267
 DB 61 TCAGAGTTATCATTCAGAAATTTTTCAGACATGCTGATCTGGTTTACCAAGAAA 120
 QY 268 GTTGCTGAAAACCTAGTAAATTTTACGTTGACGGCTAGTCTGACATAGTATTAGAT 327
 DB 121 GTTGCTGAAAACCTAGTAAATTTTACGTTGACGGCTAGTCTGACATAGTATTAGAT 180
 QY 328 GAAAGGCTATCGAAGCTTTAAAGAGTTCAATGAAGACGGCGCATTTGGAGCTTCAA 387
 DB 181 GAAAGGCTATCGAAGCTTTAAAGAGTTCAATGAAGACGGCGCATTTGGAGCTTCAA 240
 QY 388 CAGTTTAAAGACAGTATCTCTCATGTTTCAAGACAAAGTGCTTTTATGTGAGTC 447
 DB 241 CAGTTTAAAGACAGTATCTCTCATGTTTCAAGACAAAGTGCTTTTATGTGAGTC 300
 QY 448 ATGAAGACTTACAGGACAGAGAAAAACAGGGGACCAAGATAGACACTCTAGTAAAGA 507
 DB 301 ATGAAGACTTACAGGACAGAGAAAAACAGGGGACCAAGATAGACACTCTAGTAAAGA 360
 QY 508 CCAATGAGGCAAAATTAAGGCACTTTTGAAGAAACAGGCTACACACTTGAATGAGCT 567
 DB 361 CCAATGAGGCAAAATTAAGGCACTTTTGAAGAAACAGGCTACACACTTGAATGAGCT 420
 QY 568 ACAAGTCAGAGGAAGTATGAGAGACCACTCCAGATTCCTGTTTATTCAGGTCAAGCCT 627
 DB 421 ACAAGTCAGAGGAAGTATGAGAGACCACTCCAGATTCCTGTTTATTCAGGTCAAGCCT 480
 QY 628 TCTGTGCGCACTGAGATATTTGTGGGAGAGATCCCGAGAGATCTGTTTGAATGAGCT 687
 DB 481 TCTGTGCGCACTGAGATATTTGTGGGAGAGATCCCGAGAGATCTGTTTGAATGAGCT 540
 QY 688 GTTCCATTTATTTGAGAAAGCTGACCTATATGAGATCTTGTTTAAATGATGATCCG-- 744
 DB 541 GTTCCATTTATTTGAGAAAGCTGACCTATATGAGATCTTGTTTAAATGATGATCCGCTC 600

QY 745 -----CTGACGTGCTCAACAGAGTTATGAGTTTGTACATTTTGTACAAAGAGCA 798
 DB 601 ACTGCTTCACTAGCTGCTCAATAGAGTTATGAGTTTGTACATTTTGTACAAAGAGCA 660
 QY 799 GCACAAAGAGCTGTTTAACTGTTATTAATCAATGAATTCGTTCCGGAGAGCAATTTGT 858
 DB 661 GCTCAGAGAGCTGCTTAACCTGTTATTAATCAATGAATTCGTTCCGGAGAGCAATTTGT 720
 QY 859 GTTCGATTCAGTTGCCAACAATAGCTTTTGTGGGCTGATTCCTAAGATTAAC 918
 DB 721 GTTCGATTCAGTTGCCAACAATAGCTTTTGTGGGCTGATTCCTAAGATTAAC 780
 QY 919 AAGGAGCAATCTTGAAGAAATTTAGCAAAAGTACAGAGGCTCTACAGATGATTTTA 978
 DB 781 AAGGAGCAATCTTGAAGAAATTTAGCAAAAGTACAGAGGCTCTACAGATGATTTTA 840
 QY 979 TACCAACCACTGATGACAAAGAAAAAACAAGGCTTTTGTCTTGAATATGAAGAT 1038
 DB 841 TACCAACCACTGATGACAAAGAAAAAACAAGGCTTTTGTCTTGAATATGAAGAT 900
 QY 1039 CACAAAACGCTGCCAGGCAAGACGTAGGCTAATGATGTAAAGTCAAAAGCTGGGGA 1098
 DB 901 CACAAAACGCTGCCAGGCAAGACGTAGGCTAATGATGTAAAGTCAAAAGCTGGGGA 960
 QY 1099 AATGTTGAACTGTGAGTGAGGCTGATCTTATGAAGATCTGATCTGAAGTTATGCA 1158
 DB 961 AATGTTGAACTGTGAGTGAGGCTGATCTTATGAAGATCTGATCTGAAGTTATGCA 1020
 QY 1159 AAGGTAAAGTCTGTTTGTACGCAACTTGCCCAACAGGTTAAGAAATTTTGA 1218
 DB 1021 AAGGTAAAGTCTGTTTGTACGCAACTTGCCCAACAGGTTAAGAAATTTTGA 1080
 QY 1219 AAGTCAATTTAGTCAAGTTGGGAACTGAAACGATTAAGAACTAAAGATTATGCTTC 1278
 DB 1081 AAGTCAATTTAGTCAAGTTGGGAACTGAAACGATTAAGAACTAAAGATTATGCTTC 1140
 QY 1279 ATTCAATTTGATGAGAGATGCTGCTCAAGGCTATGAAAGAAATGATGTTAAGAC 1338
 DB 1141 ATTCAATTTGATGAGAGATGCTGCTCAAGGCTATGAAAGAAATGATGTTAAGAC 1200
 QY 1339 TTGAGAGGAGAAATATTGAATTTGTTTGTAAAGCCACGATCAGAGAGGAAAGA 1398
 DB 1201 TTGAGAGGAGAAATATTGAATTTGTTTGTAAAGCCACGATCAGAGAGGAAAGA 1260
 QY 1399 AGAAAAGCTCAGAGGCAAGCAAGAAATCAATGATGATGATTAATCTATTATAGCT 1458
 DB 1261 AGAAAAGCTCAGAGGCAAGCAAGAAATCAATGATGATGATTAATCTATTATAGCT 1320
 QY 1459 CCACTCATATGCTCTCCCAACAAAGAGTGAAGGCGTGAAGTGAAGTGGCTATGGA 1518
 DB 1321 CCACTCATATGCTCTCCCAACAAAGAGTGAAGGCGTGAAGTGAAGTGGCTATGGA 1380
 QY 1519 TATCCCTCAGATTTATGATGATGAAGATTTATGATTAATGATTAATGATTAACAT 1578
 DB 1381 TATCCCTCAGATTTATGATGATGAAGATTTATGATTAATGATTAATGATTAACAT 1440
 QY 1579 AACTATCGTGTGATGATGAAGATCAATGATGATGATTAATGATTAATGATTAACAT 1638
 DB 1441 AACTATCGTGTGATGATGAAGATCAATGATGATGATTAATGATTAATGATTAACAT 1500
 QY 1639 AGAGGAAGGAGTGTGAGAGAGCAAGGAGTGTCTCATCAAGAGTGTGGGCTGCT 1698
 DB 1501 AGAGGAAGGAGTGTGAGAGAGCAAGGAGTGTCTCATCAAGAGTGTGGGCTGCT 1560
 QY 1699 CCTCCCGTGTGAGAGCGGTTATTCAGAGAGAGCGCTGATCAGCAAGAGCGCTT 1758
 DB 1561 CCTCCCGTGTGAGAGCGGTTATTCAGAGAGAGCGCTGATCAGCAAGAGCGCTT 1620
 QY 1759 CGCGGTGAGAGAGAGTGCACCAACAAAGAGCGCGGAGGAGAAAGGCTGACAGGCC 1818
 DB 1621 CGCGGTGAGAGAGAGTGCACCAACAAAGAGCGCGGAGGAGAAAGGCTGACAGGCC 1677

OY	1534	TATGATATCGAAGATTATATATGATTATTAATGGTTAAATTACCAATACATCTGTGGTGA	1593
Db	1381	TATGATATGAAAGATTATTAATGATTATTAATGGTTAAATTACCAATACATCTGTGGTGA	1440
OY	1594	TATGAAGATCCACTATGATTATAGAATTTTCAGTTTGAGCTAGAGAAAGGGTGGT	1653
Db	1441	TATGAAGATCCATACTATGTTATATGAAGTTTTCAGTTTGAGCTAGAGAAAGGGTGGT	1500
OY	1654	AGAGAGCAGAGGGGCTGCTCATTCACAAGGTCGTGGGGCTGCTCTCCCCGTGGTAGA	1713
Db	1501	AGAGAGCAGAGGGGCTGCTCATTCACAAGGTCGTGGGGCTGCTCTCCCCGTGGTAGA	1560
OY	1714	GCCGGTTATTCACAGAGAGGAGGCCCTGGATCAGCAAAGGGGGTTCGGCTGGCAGAGA	1773
Db	1561	GCCGGTTATTCACAGAGAGGAGGCTCTGGATCAGCAAAGGGGGTTCGGCTGGCAGAGA	1620
OY	1774	GGTGCCCACAAACAAGAGGCCCGGGGGAAAAGGGGCTGAGAGGCCGCTCTGTA	1833
Db	1621	GGTGCCCACAAACAAGAGGCCCGGGGGAAAAGGGGCTGAGAGGCCGCTCTGTA	1680
OY	1834	CAAT 1837 	
Db	1681	CAAT 1684	
<hr/>			
RESULT 12			
LOCUS	AY034481	1872 bp	mRNA linear PRI 03-OCT-2001
DEFINITION	Homo sapiens hnRNP Q3 mRNA, complete cds.		
ACCESSION	AY034481		
VERSION	AY034481.1 GI:15809585		
KEYWORDS	"		
SOURCE	Homo sapiens. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1872) Mourtelatos,Z., Abel,L., Yong,J., Kataoka,N. and Dreyfuss,G. SMN Interacts with a novel family of hnRNP and spliceosomal proteins EMBO J. 20 (19), 5443-5452 (2001)		
JOURNAL MEDLINE PUBMED	21458434 11574476		
REFERENCE	2 (bases 1 to 1872) Mourtelatos,Z. Direct Submissions Submitted (11-MAY-2001) HMMI / Biochemistry, University of Pennsylvania, 330 CMB, 415 Curie Blvd., Philadelphia, PA 19104, USA		
AUTHORS TITLE JOURNAL FEATURES			
source	1..1872 /organism="Homo sapiens" /db_xref="taxon:9606" 1..1872 /codon_start=1 /product="hnRNP Q3" /protein_id="AAK59703.1" /db_xref="GI:15809586" /translation="WATEHVNGSGTEBPMDTSAVHSENFQTLIDGLPKVAKLID RTVAIVAHSDLEDERALELKERNEGCALVALDOPDSLSHTQNSAPFCGMKTYT RQRKKQTKYADSSKGPDEBAKIKALIBETGTTLLVTVGQRKITGPPDVSVSQQPSV GTELIFVGKIPDLPEDELVLPERKAGPLWDLRLLMDPLTGNRYAVATVECTKEAOBE AVUKYNHNIEISGRKHIGVCISVANRRLPVGSIPSKSKEQILIEFSKTELGELIVLY HQPDDKKNRSPFCELEYEDHKYTAOARRLMSGKAKWGANGVTEMADP IEDPPEVMWA AKVAVLFVRNLANTVTEBILLEKASPFKEIRVKKLADYAFIHDERDGAVKAMENNKA KGLGEMNIELVPAKPDPDKERKKAQAOKNMOMDYAYYGSPHPMPPTPRRGGRGAG RGCGYCPDDTYGYEDTDYDYTVTHNTFGCYEDFTYGEDFQVARKGGKARGAGAAP SRKGAAEPKRRAIGTSQRGGESAGVNGARGAAQQQRGRGVARGARGGVGKRR ADNGNQPSKRSQNNQNWGSPPLAODPLQGDSHGNYGVYSENOEFYDTFGQOWT"		
CDS			
BASE COUNT	590 a 323 c 489 g 470 t		
ORIGIN			
Query Match	45.0%; Score 1552.2; DB 9; Length 1872;		
Best Local Similarity	95.6%; Pred. No. 0;		

[illegible]

QY 1234 TTGGGAACTGGAACGATTAAGAGCTAAAGATTATGCTTCATTCATTTGATGAG 1293
 DB 1081 TTGGGAACTGGAACGATTAAGAGCTTAAGATTATGCTTCATTCATTTGATGAG 1140
 QY 1294 AGAGATGCTGCTCAAGGCTATGGAAGAAATGAATGCTTAAGACTTGAAGGAGAAAT 1353
 DB 1141 CGAGATGCTGCTCAAGGCTATGGAAGAAATGAATGCTTAAGACTTGAAGGAGAAAT 1200
 QY 1354 ATTGAATTTGTTTGTCTAAGCCCAAGATCAAGAGGAAAGAAAGAAAGCTCAAGG 1413
 DB 1201 ATTGAATTTGTTTGTCTAAGCCCAAGATCAAGAGGAAAGAAAGAAAGCTCAAGG 1260
 QY 1414 CAAGCAGCAAGAGATCAATATGATGATGATTAAGTACTTATGCTCAACCTCATGCT 1473
 DB 1261 CAAGCAGCAAGAGATCAATATGATGATGATTAAGTACTTATGCTCAACCTCATGCT 1320
 QY 1474 CCCCACCAAGAGGCTGAGGCGTGAAGTGAAGGCTGATATGCTTCCAGATTAAT 1533
 DB 1321 CTTCCACCAAGAGGCTGAGGCGTGAAGTGAAGGCTGATATGCTTCCAGATTAAT 1380
 QY 1534 TATGATATCAAGATTTATGATTAATGCTTATGATTAATCAATATCTATGCTGGA 1593
 DB 1381 TATGATATCAAGATTTATGATTAATGCTTATGATTAATCAATATCTATGCTGGA 1440
 QY 1594 TATGATATCAATCTATGCTTATGATTAAGTGAAGTGAAGGCTGATATGCTTCCAGATTAAT 1653
 DB 1441 TATGATATCAATCTATGCTTATGATTAAGTGAAGTGAAGGCTGATATGCTTCCAGATTAAT 1500
 QY 1654 AGAGAGCAAGGCGTGTCTCATCCAGAGTGTGCGGCTGCTCCCTCCGCTGATGA 1713
 DB 1501 AGAGAGCAAGGCGTGTCTCATCCAGAGTGTGCGGCTGCTCCCTCCGCTGATGA 1560
 QY 1714 GCGGTTATTCACAGAGAGGCGCTGATCAAGAGGCGCTTCCGCGTGCAGAGGA 1773
 DB 1561 GCGGTTATTCACAGAGAGGCGCTGATCAAGAGGCGCTTCCGCGTGCAGAGGA 1620
 QY 1774 GGTGCCCAACAAAGAGCGCGGAGGAAAGGGGTCAGGCGCGGTC 1822
 DB 1621 GGTGCCCAACAAAGAGCGCGGAGGTCAGTGTGCGAGGCGGTGCC 1669

RESULT 13
 AF408434 2232 bp mRNA linear ROD 10-OCT-2001
 LOCUS
 DEFINITION Mus musculus RRM RNA binding protein NSAP1 (Nsap1) mRNA, partial cds.
 ACCESSION AF408434 GI:15986734
 VERSION AF408434.1
 KEYWORDS Mus musculus.
 SOURCE
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2232)
 Zhou, M. and Raschke, W. C.
 Mus musculus RRM RNA binding protein, NSAP1
 Unpublished
 2 (bases 1 to 2232)
 Zhou, M. and Raschke, W. C.
 Direct Submission
 Submitted (13-AUG-2001) Immunology, Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121, USA
 JOURNAL
 FEATURES
 Source
 1. 2232
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 <1..2232
 /gene="Nsap1"
 <1..1679
 /gene="Nsap1"
 /codon_start=3
 /product="RRM RNA binding protein NSAP1"

/protein_id="AA11726.1"
 /db_xref="GI:15986735"
 /translation="KERBDEGALAVLQGFQSDSHVONKSAFLCQVKKYRQREKQ
 TKVADSSKGPDEAKIKALLERTGYTLDTTQQRKGGPPPSVSGQDPSTGTELVG
 KIPDLFEDIEVLPEKAGPIWDLNMDPLFNRRGAPVFTCTKEAQAQSVKLYNN
 HEIRSGKHIVGCI SVANNRLFVGSIPKSKTEQLIEESKTYEGLTDLVLYHOPDK
 KNRGCFLEVEDHKTAAQARRRLMSGKTKWGVNVTVMADPIEDPDENAKVYLF
 VYNNLANTYTBELKESFSPQFKLERVKLKNYAFTHPRBDGAVVAMEENNGKLEGE
 NIEIVFAKPPDQKRERAKORAKNMDYDYTYGPPHMPPTRRGRGRGTYGP
 EDYVGEDYDYDYGDYNNYRGEIEDPYTYGEDFVGARGRGRGARAPSRKGLA
 PPRGAGYSORGGPSGAVRGARGAQOQGRGVGARGRGGRGVGKRKADGVNQP
 DTKRQTNQWNGSPPIAQOPLQGGDHSNGNYKSENDEFYDFEGQWKK"
 BASE COUNT 689 a 387 c 551 g 605 t
 ORIGIN
 Query Match 42.2%; Score 1456.8; DB 10; Length 2232;
 Best Local Similarity 99.2%; Pred. No. 8-5e-305;
 Matches 1464; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 347 TAAAGAGTTCATGTAAGACGGCGCATTTGGCAGTCTTCAACAGTTTAAAGACATGATC 406
 DB 1 TAAAGAGTTCATGTAAGACGGCGCATTTGGCAGTCTTCAACAGTTTAAAGACATGATC 60
 QY 407 TCTCTATGTCAGAACAAAGAGCTTTTATGCTGATGAAAGCTTACAGCAGA 466
 DB 61 TCTCTATGTCAGAACAAAGAGCTTTTATGCTGATGAAAGCTTACAGCAGA 120
 QY 467 GAGAAAACAGGGGACCAAGATGACAGACTGTATGAAGACCAAGTGAAGGCAAGATTA 526
 DB 121 GAGAAAACAGGGGACCAAGATGACAGACTGTATGAAGACCAAGTGAAGGCAAGATTA 180
 QY 527 AGGACCTTTGAAAACAGAGCTACACCTGATGCTACAGCTGACAGAGAGTATG 586
 DB 181 AGGACCTTTGAAAACAGAGCTACACCTGATGCTACAGCTGACAGAGAGTATG 240
 QY 587 GAGAACCACTCCAGATTCGTTTATTCAGTCAAGAGCTTCTGTTGGCACTGAGATAT 646
 DB 241 GAGAACCACTCCAGATTCGTTTATTCAGTCAAGAGCTTCTGTTGGCACTGAGATAT 300
 QY 647 TTGTGGGAAAGATCCCAAGAGATCTGTTTGAAGATGAGCTTGTCCATTTATGAGAAAG 706
 DB 301 TTGTGGGAAAGATCCCAAGAGATCTGTTTGAAGATGAGCTTGTCCATTTATGAGAAAG 360
 QY 707 CTGGACCTTATGAGGATCTTCGTTTATGATGATGATCCGCTCATCTGCTCAACAGAGTT 766
 DB 361 CTGGACCTTATGAGGATCTTCGTTTATGATGATGATCCGCTCATCTGCTCAACAGAGTT 420
 QY 767 ATGCGTTTGTCACTTTTGTACAAAGAGACAGCAAGAGCTGTTAACTGTATATA 826
 DB 421 ATGCGTTTGTCACTTTTGTACAAAGAGACAGCAAGAGCTGTTAACTGTATATA 480
 QY 827 ATCATGAATTTGTTCCGGGAGACACATTTGCTGATCTTCAAGTTCACCAATATGAGC 886
 DB 481 ATCATGAATTTGTTCCGGGAGACACATTTGCTGATCTTCAAGTTCACCAATATGAGC 540
 QY 887 TTTTGTGGGCTCGATTTCTTAAGAGTAAACCAAGAGAGATTTCTTGAAGATTTTGA 946
 DB 541 TTTTGTGGGCTCGATTTCTTAAGAGTAAACCAAGAGAGATTTCTTGAAGATTTTGA 600
 QY 947 AAGTGACAGAGGCTCACAGATGTCATTTTATCACCAACCTGATGACAGAAAAA 1006
 DB 601 AAGTGACAGAGGCTCACAGATGTCATTTTATCACCAACCTGATGACAGAAAAA 660
 QY 1007 ACAGAGGCTTTGCTTCTTGAATATGAAATCAAAAACAGCTGCCAGGCAAGCGTA 1066
 DB 661 ACAGAGGCTTTGCTTCTTGAATATGAAATCAAAAACAGCTGCCAGGCAAGCGTA 720
 QY 1067 GCGTATATGATGTAAGTCAAGTCTGGGGAATTTTGAATCTGTTGATGGCTGATC 1126
 DB 721 GCGTATATGATGTAAGTCAAGTCTGGGGAATTTTGAATCTGTTGATGGCTGATC 780
 QY 1127 CATTGAATCTCTGATCTGAAGTATGCAAGGTAAGGTAAGGCTGTTGATGCAACC 1186

```

Db      781  CTAATGAGATCCTGATCTGAGTTATGCAAGGTAAGGTAAGGTCGTGTTGTACGCAACC 840
Oy      1187 TTGCAACACGATACGAGAAATTTAGAAAGTCACTTAGCTTTGGGAACTGG 1246
Db      841  TTGCAACACGATACGAGAAATTTAGAAAGTCACTTAGCTTTGGGAACTGG 900
Oy      1247 AACGAGTAAAGAGTAAAGATTTATGCTTTCACTTTTGTATGAGAGATGTCG 1306
Db      901  AACGAGTAAAGAGTAAAGATTTATGCTTTCACTTTTGTATGAGAGATGTCG 960
Oy      1307 TCAGGCTATGAGAAATGATGTAAGACTTGAGGAGGAAATATTGAAATGTTT 1366
Db      961  TCAGGCTATGAGAAATGATGTAAGACTTGAGGAGGAAATATTGAAATGTTT 1020
Oy      1367 TTGCTAAGCCACGATCAGAGAGGAAAGAAAGAAAGCTCAGAGCAAGCAAGCA 1426
Db      1021 TTGCTAAGCCACGATCAGAGAGGAAAGAAAGAAAGCTCAGAGCAAGCAAGCA 1080
Oy      1427 ATCAAAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1486
Db      1081 ATCAAAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Oy      1487 GTGAGAGCGGTGAGAGTGAAGGTGCTATGATGATGATGATGATGATGATGATGATG 1546
Db      1141 GTGAGAGCGGTGAGAGTGAAGGTGCTATGATGATGATGATGATGATGATGATGATG 1200
Oy      1547 ATATATATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1606
Db      1201 ATATATATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
Oy      1607 ACTATGATTAATGAAGATTTTCAAGTTGAGAGTGAAGAGGAGGAGTGAAGAGCAAGG 1666
Db      1261 ACTATGATTAATGAAGATTTTCAAGTTGAGAGTGAAGAGGAGGAGTGAAGAGCAAGG 1320
Oy      1667 GTGCTGCTCCTCATCAGAGAGTGTGAGGAGTGTCTCTCTCCCGTGTGAGAGCCGTTATTCAC 1726
Db      1321 GTGCTGCTCCTCATCAGAGAGTGTGAGGAGTGTCTCTCTCCCGTGTGAGAGCCGTTATTCAC 1380
Oy      1727 AAGAGAGAGGCGCTGATCAGAGAGGCGTTGGCGGTGAGAGAGGAGTGTGAGAGCAAC 1786
Db      1381 AAGAGAGAGGCGCTGATCAGAGAGGCGTTGGCGGTGAGAGAGGAGTGTGAGAGCAAC 1440
Oy      1787 AAAGAGGCGCGGAGGAAAGAGGAGTGTGAGAGGCGGATC 1822
Db      1441 AAAGAGGCGCGGAGGTAAGTGTGAGAGGAGTGTGAGG 1476

RESULT 14
AL136082 74337 bp DNA linear RPI 18-JUL-2001
LOCUS AL136082/c
DEFINITION Human DNA sequence from clone RPI-3J17 on chromosome eq14.3-16.2,
complete sequence.
ACCESSION AL136082
VERSION AL136082.22 GI:14970778
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 74337)
Requester: clonerequest@sanger.ac.uk
On Jul 19, 2001 this sequence version replaced gi:14800017.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all

```

FEATURES

source

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats, all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RPI-3J17 is from the library RPI-1 constructed by the group of Pletzer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCVPAC2

IMPORTANT: This sequence is not the entire insert of clone RPI-3J17 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RPI-3J17 is at 1 in this sequence. The true left end of clone RPI-33B24 is at 72338 in this sequence. The true right end of clone RPI-321M4 is at 55916 in this sequence.

Location/Qualifiers

1..74337	Organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
	/map="q14.3-16.2"
	/clone="RPI-3J17"
	/clone_lib="RPI-1"
5..786	
/note="L1PA7 repeat: matches 5346..6141 of consensus"	
825..995	
/note="L1M4 repeat: matches 3047..3242 of consensus"	
972..1237	
/note="L1MCS repeat: matches 6956..7221 of consensus"	
1396..1589	
/note="L1MCS repeat: matches 7299..7485 of consensus"	
1590..1778	
/note="MER58C repeat: matches 1..88 of consensus"	
1779..1970	
/note="L1MCS repeat: matches 7485..7699 of consensus"	
1975..2181	
/note="AluIo repeat: matches 132..302 of consensus"	
2182..2459	
/note="AluSk repeat: matches 26..294 of consensus"	
2460..2577	
/note="AluIo repeat: matches 5..132 of consensus"	
2613..2721	
/note="L1MCS repeat: matches 7813..7929 of consensus"	
2729..2877	
/note="L1M4 repeat: matches 3229..3373 of consensus"	
2878..3163	
/note="AluIb repeat: matches 1..289 of consensus"	
3164..3409	
/note="L1M4 repeat: matches 3373..3612 of consensus"	
3410..3690	
/note="AluSk repeat: matches 1..293 of consensus"	
3694..4001	
/note="AluY repeat: matches 1..305 of consensus"	
4002..4046	
/note="L1M4 repeat: matches 3612..3684 of consensus"	
4054..4627	
/note="TIGBR2 repeat: matches 2098..2677 of consensus"	
4628..4919	
/note="AluIb repeat: matches 1..303 of consensus"	
4920..6450	
/note="TIGBR2 repeat: matches 597..2098 of consensus"	
6453..6515	
/note="TIGBR2 repeat: matches 2654..2717 of consensus"	

```
repeat_region 7130..7437
/note="AluX repeat: matches 1..305 of consensus"
repeat_region 8110..8427
/note="AluY repeat: matches 1..293 of consensus"
misc_feature 8522..8983
/note="match: GSS: Em:AQ232795"
misc_feature 8543..8940
/note="match: GSS: Em:B58596"
misc_feature 8798..9745
/note="CpG Island"
/evidence=not experimental
misc_feature /complement(9078..9313)
/note="match: GSS: Em:AQ938728"
misc_feature 9330..9679
/note="match: GSS: Em:AQ938963"
repeat_region 11798..11883
/note="MER3 repeat: matches 36..2341 of consensus"
repeat_region 11998..12071
/note="MER3 repeat: matches 36..2341 of consensus"
repeat_region 12243..12373
/note="MER3 repeat: matches 2505..2585 of consensus"
misc_feature /note="L2 repeat: matches 2616..2749 of consensus"
/complement(14486..14660)
/note="match: STR: Em:G44117"
repeat_region 14713..14771
/note="AluY repeat: matches 86..144 of consensus"
repeat_region 14772..15071
/note="AluX repeat: matches 1..300 of consensus"
repeat_region 15072..15224
/note="AluY repeat: matches 143..298 of consensus"
repeat_region 15302..15611
/note="AluX repeat: matches 1..312 of consensus"
repeat_region 15876..15918
/note="MER3 repeat: matches 2528..2570 of consensus"
repeat_region 16462..16726
/note="AluX repeat: matches 1..268 of consensus"
repeat_region 17470..17495
/note="AluX repeat: matches 1..268 of consensus"
repeat_region 17999..18286
/note="13 copies 2 mer aa 100% conserved"
repeat_region 18734..18866
/note="AluX repeat: matches 1..305 of consensus"
repeat_region 18867..19165
/note="AluY repeat: matches 1..131 of consensus"
repeat_region 19166..19329
/note="AluX repeat: matches 1..297 of consensus"
repeat_region 20415..20556
/note="AluY repeat: matches 131..302 of consensus"
repeat_region 20557..20837
/note="FRAM repeat: matches 2..131 of consensus"
misc_feature /note="AluX repeat: matches 1..302 of consensus"
20625
/note="Weak data"
repeat_region 20838..20867
/note="AluY/FRAM repeat: matches 267..298 of consensus"
repeat_region 21848..21922
/note="L2 repeat: matches 2580..2656 of consensus"
misc_feature /complement(23637..24080)
/note="match: GSS: Em:AQ075660"
misc_feature /complement(23885..24499)
/note="match: GSS: Em:AQ976004"
misc_feature 24510..24905
/note="match: GSS: Em:AQ0703263"
join(28254..28325,30383..30749,34220..34344)
/note="match: GSS: Em:AQ008463"
repeat_region 31066..31220
/note="AluX/x repeat: matches 137..291 of consensus"
repeat_region 31726..32049
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 32199..32402
/note="MER20 repeat: matches 2..218 of consensus"
repeat_region 32403..32585
/note="MIR repeat: matches 9..202 of consensus"
repeat_region 33032..33082
/note="MIR repeat: matches 91..144 of consensus"
```

```
repeat_region 33661..33963
/note="AluX repeat: matches 1..301 of consensus"
repeat_region 34949..35007
/note="L1MC/D repeat: matches 5609..5668 of consensus"
repeat_region 35003..35367
/note="L1MC1 repeat: matches 5329..5566 of consensus"
repeat_region 35368..35665
/note="AluX repeat: matches 2..300 of consensus"
repeat_region 35666..36207
/note="L1MC1 repeat: matches 5656..6326 of consensus"
repeat_region 36738..37037
/note="AluY repeat: matches 5..310 of consensus"
repeat_region 37493..37791
/note="AluX repeat: matches 1..294 of consensus"
repeat_region 38684..38985
/note="AluX repeat: matches 1..304 of consensus"
repeat_region 39121..39343
/note="MIR repeat: matches 20..260 of consensus"
misc_feature 39257..39521
/note="match: GSS: Em:A2610695"
repeat_region 40783..41087
/note="AluY repeat: matches 1..302 of consensus"
repeat_region 41255..41461
/note="MER20 repeat: matches 1..218 of consensus"
repeat_region 42386..42690
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 43456..43757
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 44850..44943
/note="AluX repeat: matches 43..136 of consensus"
```

Query Match 39.3%; Score 1356.4; DB 9; Length 74337;
Best Local Similarity 96.0%; Pred.No. 6; 7e-283;
Matches 1545; Conservative 0; Mismatches 41; Indels 24; Gaps 14;

```
QY 1800 GGGAAAAGGGGTGAGCGCGCTGACCTGTATGATGAGATGATGCTGATTTGTG 1859
28323 GGGAAAAGGGGTGAGCGCGCTGACCTGTATGATGAGATGATGCTGATTTGTG 28265
QY 1860 GATTACACGAGAGCTTGACGTGAGTAAATGTAAGAAATCAACAACTTAAATATC 1919
28264 GATTACACGAGAGCTTGACGTGAGTAAATGTAAGAAATCAACAACTTAAATATG 28206
QY 1920 TCGGATGTATGAGCATATCTATGCGAAGACCTCTATGAGATGAGATGCA 1979
28205 TCGGATGTATGAGCATATCTATGCGAAGACCTCTATGAGATGAGATGCA 28146
QY 1980 ATACGGGACATGACATTAATCTTGACCTTGTATGATTTCTTAAATTTCTG 2039
28145 ATACGGGACATGACATTAATCTTGACCTTGTATGATTTCTTAAATTTCTG 28086
QY 2040 CAGTGCAGTTATTAATCAATTAAGCTATCTATTTCCAAATGTTCCAA-AAAAATCCT 2098
28085 CAGTGCAGTTATTAATCAATTAAGCTATCTATTTCCAAATGTTCCAAAGAAATCCT 28026
QY 2099 TCAATTAATCTTACATGATGATCTATTAAGAAATTAAGTTGTTCTTTAAAAATCTGC 2158
28025 TCAATTAATCTTACATGATGATCTATTAAGAAATTAAGTTGTTCTTTAAAAATCTGC 27971
QY 2159 TCAATTAATCTTACATGATGATCTATTAAGAAATCAAGATGATGATCTGAAATA 2218
27970 TCAATTAATCTTACATGATGATCTATTAAGAAATCAAGATGATGATCTGAAATA 27912
QY 2219 TTCTCTGAATTTGTCATTAATTAATTTATGAGATGATGAG-ATGCCATGTTGG 2274
27911 TTCTCTGAATTTGTCATTAATTAATTTATGAGATGATGATGAGATGAGATGAG 27852
QY 2275 TACCTTAATTTATTTCTGCTACCAAGTTATATCATGATGATCTATTTTATTA 2234
27851 TACCTTAATTTATTTCTGCTACCAAGTTATATCATGATGATCTATTTTATTA 27793
QY 2335 TATTAATCACTTTGATTTGTTGAG-ATGAGTTTCAAGTATATCATGAGATGAG 2392
```

Db 27792 TAGTATCACTTTGTAATGTGTACAGATATGCAATTCAGGTGTAATCATCAGAGCTGG 27733
 Qy 2393 TTATGACGACATTCAGATATGAGTGTCTTTTCAGAACCTTTT-AAAGGGTGGTAACT 2451
 Db 27732 TTATGACGACATTCAGATATGAGTGTCTTTTCAGAACCTTTT-AAAGGGTGGTAACT 27673
 Qy 2452 ACCTGAGTACAGAGAGTGAATTAACCTGTCTGTACTGTATCATAGAAATCTTTGTA 2511
 Db 27672 ACCTGAGTACAGAGAGTGAATTAACCTGTCTGTACTGTATCATAGAAATCTTTGTA 27613
 Qy 2512 GATPAAAGCAAGGCTGTAAATATGATATGAGGGTAAAGTTTAAATATACCAATGTAA 2571
 Db 27612 GATPAAAGCAAGGCTGTAAATATGATATGAGGGTAAAGTTTAAATATACCAATGTAA 27553
 Qy 2572 CATCTTAGTGTGCTTAGTTTTCAGAGGCTGTAGAGCTTCCATGATCAGATCATTAACAG 2631
 Db 27552 CATCTTAGTGTGCTTAGTTTTCAGAGGCTGTAGAGCTTCCATGATCAGATCATTAACAG 27493
 Qy 2632 GCGTTGCTTTTGTGTAATTTTGTGGCTGMAAAAGCAGCCTTGCTTTCAGATATTTAG 2691
 Db 27492 GCGTTGCTTTTGTGTAATTTTGTGGCTGMAAAAGCAGCCTTGCTTTCAGATATTTAG 27433
 Qy 2692 TTATTTGATATATATATGTTTTCAGAAAGTGTACTTTTGTAAAGCATCATGATGTTCAA 2751
 Db 27432 TTATTTGATATATATATGTTTTCAGAAAGTGTACTTTTGTAAAGCATCATGATGTTCAA 27376
 Qy 2752 AAAAAAGTGCATCCGAACCTGTACTTAAATCTGCAAGTCCCTTAAATAAAGTCAGACT 2811
 Db 27375 AAAAAAGTGCATCCGAACCTGTACTTAAATCTGCAAGTCCCTTAAATAAAGTCAGACT 27316
 Qy 2812 AAAAGTGCATATGTATCAGCAAGGCTGTACATTTGATTTTGAAGTTTTCATTAAT 2871
 Db 27315 AAAAGTGCATATGTATCAGCAAGGCTGTACATTTGATTTTGAAGTTTTCATTAAT 27256
 Qy 2872 CATAGAAATATGATATGCTGTATGTTTAACTTTTAAAGGTATGTTTCAATAGT 2931
 Db 27255 CATAGAAATATGATATGCTGTATGTTTAACTTTTAAAGGTATGTTTCAATAGT 27196
 Qy 2932 GCATTTGTTATGCTGATCATTAATAAATGTAATCAGTTTCCATTTCTTATGACAGGT 2991
 Db 27195 GCATTTGTTATGCTGATCATTAATAAATGTAATCAGTTTCCATTTCTTATGACAGGT 27136
 Qy 2992 CATGATTAATGTAATATAGTATCATATCTTTGCTATGTTTAAATTTCTAAAGCA 3051
 Db 27135 CATGATTAATGTAATATAGTATCATATCTTTGCTATGTTTAAATTTCTAAAGCA 27076
 Qy 3052 CCTTATGATGATGATGATGCTGATGCTGTAAGCATCTCTATATGATCACTCAAAAAT 3111
 Db 27075 CCTTATGATGATGATGATGCTGATGCTGTAAGCATCTCTATATGATCACTCAAAAAT 27017
 Qy 3112 TTTTGTGCAAGCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3170
 Db 27016 TTTTGTGCAAGCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26957
 Qy 3171 GGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3229
 Db 26956 GGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26897
 Qy 3230 AACCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3288
 Db 26896 AACCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26837
 Qy 3289 CCAAGTCCCAAGTAAATGCAATTTTAAAGAAATTTAAAGCAATGATGATGATGATGATGATGAT 3348
 Db 26836 CCAAGTCCCAAGTAAATGCAATTTTAAAGAAATTTAAAGCAATGATGATGATGATGATGATGAT 26777
 Qy 3349 CTCATGCAAAACATATTTGATTCCTGCAATATTAATTTGTTTATAC 3398
 Db 26776 CTCATGCAAAACATATTTGATTCCTGCAATATTAATTTGTTTATAC 26727

RESULT 15
 AY034482

LOCUS AY034482 1767 bp mRNA linear PRI 03-OCT-2001
 DEFINITION Homo sapiens hnRNP Q2 mRNA, complete cds.
 ACCESSION AY034482
 VERSION AY034482.1 GI:15809587
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1767)
 AUTHORS Mourelatos, Z., Abel, L., Yong, J., Kataoka, N. and Dreyfuss, G.
 TITLE SMN interacts with a novel family of hnRNP and spliceosomal proteins
 JOURNAL EMBO J. 20 (19), 5443-5452 (2001)
 MEDLINE 21458434
 PUBMED 11574476
 REFERENCE 2 (bases 1 to 1767)
 AUTHORS Mourelatos, Z.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2001) HHMI/ Biochemistry, University of Pennsylvania, 330 CBH, 415 Curie Blvd., Philadelphia, PA 19104, USA
 FEATURES
 source location/Qualifiers
 1..1767
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..1767
 /codon_start=1
 /product="hnRNP Q2"
 /protein_id="AAK59704.1"
 /db_xref="GI:15809588"
 translation="MATEHVNNGNEEPMOTTSVISHSENFOTLLDAGLPQVAKELD EIVYAGLVASDLDERRAIEALKERNEGALVLOQFKDSHVONKSAFLCGVNTY KOREKQGTKVADSSGPDPEAKIKALEETGATLDTTGQRKRGPPSPVSGQPSV GTELPVNGKIPDLPEDELVLPFEKAGPIWDLRLMDPLTGAGVAFTECTKEAOR AVKIVNNHISGKIIQVCI SVANNRLPVGSIIPSKTSEGLLEBSKVEGLTDLVILY HOPDCKNRSFCLEVEDHTAAQVYLFRNLANTYTESILKEAASQPKLEEVKK LKDYAFHFDNRDGAIVKMEEMNGDLEGNEIETVPAPRPOKREKRAQKQALRNQM YDDTYTPPHMPTTRRGKRGKRGCTGTPPDYGYEDYDYIYTGIDYNNYGEDPY YGEDPVQAGARGRGARGARGAAPSRGRGAAPRGAAGVSGRGPSAGPSARGAGAO QORGVGARGARGRGARGRGARADKDVYQNPDSKRRQTNWNGSQPIAQDPLQGDHS GNYGKSNSEFYQDPFGQWKK"
 BASE COUNT 561 a 310 c 453 g 443 t
 ORIGIN
 Query Match 38.9%; Score 1343.4; DB 9; Length 1767;
 Best Local Similarity 89.8%; Pred. No. 3e-280;
 Matches 1498; Conservative 0; Mismatches 66; Indels 105; Gaps 1;
 Qy 154 ATGGCTACAGAACATGTTAATGGAATGCTACTGAAAGCCCATGATATCTTCAACA 213
 Db 1 ATGGCTACAGAACATGTTAATGGAATGCTACTGAAAGCCCATGATATCTTCTGCA 60
 Qy 214 GTTATCATTTGAGAAATTTTTCAGACATTCCTGATGCTGTTTCCACAGAAAGTGGT 273
 Db 61 GTTATCATTTGAGAAATTTTTCAGACATTCCTGATGCTGTTTCCACAGAAAGTGGT 120
 Qy 274 GAAAACTAGATGAATTTTACGTTCAAGGCTAGTTGACATATGATGATTTAGATGAAGA 333
 Db 121 GAAAACTAGATGAATTTTACGTTCAAGGCTAGTTGACATATGATGATTTAGATGAAGA 180
 Qy 334 GGTATCGAAGCTTTTAAAGAGTTCAATGAGCGCGCATTTGGCAGTGTTCACAGTTT 393
 Db 181 GGTATCGAAGCTTTTAAAGAGTTCAATGAGCGCGCATTTGGCAGTGTTCACAGTTT 240
 Qy 394 AAAGCAGTATCTCTCATGTTTCAAGAACAAAGTGCCTTTTATATGAGAGTCATGAAG 453
 Db 241 AAAGCAGTATCTCTCATGTTTCAAGAACAAAGTGCCTTTTATATGAGAGTCATGAAG 300
 Qy 454 ACTTACAGGCAAGAGAAAAACAGGGGACCAAAGTAGACAGACTTATGTAAGACAGAT 513
 Db 301 ACTTACAGGCAAGAGAAAAACAGGGGACCAAAGTAGACAGATTTATGTAAGACAGAT 360
 Qy 514 GAGGCAAAAGATTAAAGCACTTTTGAAGAAAGACAGGCTACACACTTGAATGTACAGGT 573


```

Db 361 GAGGCAAAATTTAAGGACCTTTGGAAAGAACAGGCTACACACTTGAATGACACCTGGA 420
Qy 574 CAGAGAGATATGAGAGACCACTCCAGATTCGGTTTATTCAGTCCAGACCTTCGT 633
Db 421 CAGAGAGATATGAGAGACCACTCCAGATTCGGTTTATTCAGTCCAGACCTTCGT 480
Qy 634 GGCACATGAGATATTTGGGGGAGAGATCCCCAGAGATCTGTTGAGATGAGCTTTTCCA 693
Db 481 GGCACATGAGATATTTGGGGGAGAGATCCCCAGAGATCTATTTGAGATGAACTTTTCCA 540
Qy 694 TTAATTTAGAAAGCTGAGACCTATATGGAATCTTGTATATGATGATCCGCTCAGTGT 753
Db 541 TTAATTTAGAAAGCTGAGACCTATATGGAATCTTGTATATGATGATCCGCTCAGTGT 600
Qy 754 CTCACAGAGATTATGCGTTTGTCACTTTTGTATCAAAAGAGCAGACAAAGGCTGTT 813
Db 601 CTCACATGAGATTATGCGTTTGTCACTTTTGTATCAAAAGAGCAGCTCAGAGGCTGTT 660
Qy 814 AAACGTATATATATCATGAAATTCGTTCCGGAGACATTTGGTCTGCACTCAGTT 873
Db 661 AAACGTATATATATCATGAAATTCGTTCCGGAGAACATATTTGGTCTGCACTCAGTT 720
Qy 874 GCCAACATAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAAGAGCAGATTCCT 933
Db 721 GCCAACATAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAAGAGCAGATTCCT 780
Qy 934 GAGGAATTTAGCAAGATGACAGAGGCTCAAGATGTCATTTTATACCAACCACTGAT 993
Db 781 GAGGAATTTAGCAAGATGACAGAGGCTCTTACAGCGTCATTTTATACCAACCGGAT 840
Qy 994 GACAGAGAAATTAACAGAGGCTTTGCTTCTTGAATATGAAATCAAAAACAGCTGCC 1053
Db 841 GACAGAGAAATTAACAGAGGCTTTGCTTCTTGAATATGAAATCAAAAACAGCTGCC 900
Qy 1054 CAGGCAAGACGTAGCTATATAGTGTAAAGTCAAAAGCTGGGAAATGTTGAACTGTT 1113
Db 901 C----- 901
Qy 1114 GAGTGGCTGATCCTATTTGAAGATCTGATCTGAAAGTTATGGCAAGTAAGTAAGTCTG 1173
Db 902 -----AGTAAAGTCTG 915
Qy 1174 TTTGTAGCAACCTTGCCACACAGGTAACAGAGAAATTTAGAAAGTCAATTTAGTCAG 1233
Db 916 TTTGTAGCAACCTTGCCACACAGGTAACAGAGAAATTTAGAAAGTCAATTTAGTCAG 975
Qy 1234 TTTGGAAACTGAAACGATTAAGAGTAAAGATTATGCTTCATTCAATTTTGAATGAG 1293
Db 976 TTTGGAAACTGAAACGATTAAGAGTAAAGATTATGCTTCATTCAATTTTGAATGAG 1035
Qy 1294 AGAGATGCTGCTCAAGGCTATGAGAGAAATGAATGCTTAAAGCTTGAGAGAGAAAT 1353
Db 1036 CGAGATGCTGCTCAAGGCTATGAGAGAAATGAATGCTTAAAGCTTGAGAGAGAAAT 1095
Qy 1354 ATTGAATTTGTTTGTCTAAGCCACAGATCAAGAGAGAAAGAAAGAAAGCTCAGAGG 1413
Db 1096 ATTGAATTTGTTTGTCTAAGCCACAGATCAAGAGAGAAAGAAAGAAAGCTCAGAGG 1155
Qy 1414 CAACAGCAAAAGATCAAAATGATGATGATTAATTAATGCTCACTCATATGCTCT 1473
Db 1156 CAACAGCAAAAGATCAAAATGATGATGATTAATTAATGCTCACTCATATGCTCT 1215
Qy 1474 CCCCCAACAAGAGTCAAGGCGGTGAGAGTAAAGGTCATATGATATCTCCAGATTAT 1533
Db 1216 CCTCCACAAGAGGTCAGGCGGTGAGAGTAAAGGTCATATGATATCTCCAGATTAT 1275
Qy 1534 TATGATATGAAAGTATTAATGATGATGATTAATTAATGATGATTAATGATGATGAT 1593
Db 1276 TATGATATGAAAGTATTAATGATGATGATTAATTAATGATGATGATGATGATGAT 1335
Qy 1594 TATGAGATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1653

```

```

Db 1336 TATGAGATCCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
Qy 1654 AGAGAGCAAGGAGTCTGCTCATCAAGAGTCTGAGGCTGCTCTCCCGTGTAGA 1713
Db 1396 AGAGAGCAAGGAGTCTGCTCATCAAGAGTCTGAGGCTGCTCTCCCGTGTAGA 1455
Qy 1714 GCGGTTATTTCAAGAGAGAGGCGCTGATCAGCAAGAGGCGTTGCGGTCGAGAGGA 1773
Db 1456 GCGGTTATTTCAAGAGAGAGGCTGATCAGCAAGAGGCGTTGCGGTCGAGAGGA 1515
Qy 1774 GGTGCCCAACAACAAGAGGCGCGGAGGAGAAAGGAGTCAAGGCGGTC 1822
Db 1516 GGTGCCCAACAACAAGAGGCGCGGAGTACGTGTCGAGAGGAGTGGCC 1564

```

Search completed: April 28, 2003, 00:54:54
Job time : 18678.6 secs

GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 27, 2003, 16:49:41 ; Search time 8469.41 Seconds

(without alignments)
5783.170 Million cell updates/sec

Title: US-09-821-687-3_COPY_154_1836

Perfect score: 1683
Sequence: 1 atggtctacgacatcttaa.....ccgctccctgactgttaca 1683

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sbs:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_ocher:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_ocher:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	3452	10	AB035725
2	1602	95.2	2373	10	AF093821
3	1583.8	94.1	1686	9	AY034483
4	1574	93.5	2221	9	AF155568
5	1552.2	92.2	1872	9	AY034481
6	1552.2	92.2	2932	9	AF037448
7	1519	90.3	2079	6	AR177851
8	1456.8	86.6	2232	10	AF408434
9	1442.8	85.7	103819	9	HSJ775C13
10	1442.8	85.7	212659	2	AL390737
11	1366.8	81.2	3319	9	BC032643
12	1343.4	79.8	1767	9	AY034482
13	1222.8	72.7	3101	6	AX011753
14	896	53.2	1899	10	AF441128
15	895.8	53.2	151297	2	AC118904
16	875	52.0	2644	9	AF000364
17	860.8	51.1	221309	2	AC122217
18	856	50.9	2722	9	BC001449
19	853.8	50.7	2471	10	BC004679
20	816	48.5	181810	2	AC126148
21	762.4	45.3	121304	2	AC079134
22	762.4	45.3	196488	2	AC084410
23	736.4	43.8	934	6	AX013715
24	706.2	42.0	202374	10	AC122868
25	636.8	37.8	139996	9	AL161799
26	570.6	33.9	129237	2	AC127932
27	484	28.8	110783	2	AC129693
28	426	25.3	118429	2	AC119124
29	398.4	23.7	57824	2	AC101053
30	381.8	22.7	2606	3	AY058477
31	368	21.9	384	6	AR079026
32	358.2	21.3	74337	9	AL136082
33	354	21.0	2195	10	BC004001
34	351	20.9	175221	2	AC111832
35	289.4	17.2	2467	9	BC028588
36	250.8	14.9	1785	10	AY028945
37	250.8	14.9	1863	10	AF290984
38	242.4	14.4	1770	9	AF272079
39	242.4	14.4	1794	9	HSJ272078
40	242.4	14.4	1945	9	AF209192
41	240.8	14.3	1761	9	AF271789
42	223	13.8	1911	9	AF271790
43	228.6	13.6	117535	10	AL672208
44	228.6	13.6	218193	2	AC121608
45	222.8	13.2	2105	9	AK000324

ALIGNMENTS

RESULT 1
AB035725
LOCUS AB035725
DEFINITION Mus musculus SYNCRIP mRNA, complete cds.
ACCESSION AB035725
VERSION AB035725.1 GI:6576814
KEYWORDS SYNCRIP.
SOURCE Mus musculus (strain:DDV) postnatal 21 day cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (sites)
Mizutani, A., Fukuda, M., Iibata, K., Shiraishi, Y. and Mikoshiba, K.
SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

isoforms
J. Biol. Chem. 275 (13), 9823-9831 (2000)
20200483
2 (bases 1 to 3452)
Mizutani, A., Fukuda, M. and Mikoshiba, K.
Direct Submision
Submitted (11-DEC-1999) Akihiro Mizutani, Brain Science Institute,
RIKEN, the Laboratory for Developmental Neurobiology, 2-1, Hirosewa,
Wako, Saitama 351-0198, Japan (E-mail: amizutani@ims.u-tokyo.ac.jp,
Tel: +81-48-467-9745, Fax: +81-48-467-9744)
Location/Qualifiers

FEATURES
Source

1. 3452
/organism="Mus musculus"
/strain="DDY"
/db_xref="taxon:10090"
/dev_stage="postnatal 21 day"
1. 3452
/gene="SYNCRIP"
1. 3452
/gene="SYNCRIP"
/evidence=experimental
154. 1839
/gene="SYNCRIP"
/protein_id="BA88342.1"
/codon_start=1
/db_xref="GI:6576815"
/translation="MATEHNGCTEPHNTSVAHSENPQTLLDGLPQKAEKID
EIVVAGLVASDLDERALEKFNEDGALVLDQFQSDLSHVQNSATLGVMTKY
RORBKQSTKVASDSKGPDEAKIKALERTYTLDTVGQRKGGPPDSVSGQPSV
GTETPVGKIPRDLFEDELVPLFEKAGPIMDLRLMDPLQNLNGSAFVETCEAAOE
AVKLYNHEIRSGHIGVCSIVANNRPFVSIPEKSKTEGILIEFSVTEGLIDVILY
HOPDDKKNKGFCELEYEDHKTAAARRRMSGVKVMGVATVEMADPIEDDPERVM
AKVULPVRLANTVTRIELLEKSPQSGKIERKKYKDYAFHPRDEDEGAKAMEEN
GKIDENITIVPAKPPDQKKEKQKQAKOMDYDYTGPPHPPPTRRGKGG
RGSGYPPDYGYEDYDYGYDYNHNGRYEDPYGYEDPYGVGARGRGRGARAP
SRGRGAPPRGRGYSQRRGSGRGRGVRGARGAQQQRGRGKGVKVEGPDLLQ"

BASE COUNT 1056 a 623 c 777 g 996 t
ORIGIN

Query Match 100.0%; Score 1683; DB 10; Length 3452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTACAGAACTGTATATGAAATGGTACTGAAGCCCATGAGTACTTACGCA 60
DB 154 ATGGCTACAGAACTGTATATGAAATGGTACTGAAGCCCATGAGTACTTACGCA 213
QY 61 GTTATCATTCAGAAATTTTCAACATTCGTTGATGCTGTTTACACAGAAAGTTGCT 120
DB 214 GTTATCATTCAGAAATTTTCAACATTCGTTGATGCTGTTTACACAGAAAGTTGCT 273
QY 121 GAAATACTAGTAAATTTAGTTGACAGGCTAGTTCACATAGTATTTAGTAAAGA 180
DB 274 GAAATACTAGTAAATTTAGTTGACAGGCTAGTTCACATAGTATTTAGTAAAGA 333
QY 181 GCTATCGAAGCTTTAAAGAGTTCAATGAGCGCGCATTTGGAGTCTCAACAGTTT 240
DB 334 GCTATCGAAGCTTTAAAGAGTTCAATGAGCGCGCATTTGGAGTCTCAACAGTTT 393
QY 241 AAAGACAGTATCTCTCATGTTCAAGCAAAAGTGCCTTTTATGTGAGTCAATGAG 300
DB 394 AAAGACAGTATCTCTCATGTTCAAGCAAAAGTGCCTTTTATGTGAGTCAATGAG 453
QY 301 ACTTACAGCGAGAGAGAAAAAGAGGACCAAGTAGACACTCTAGTAAAGACAGAT 360
DB 454 ACTTACAGCGAGAGAGAAAAAGAGGACCAAGTAGACACTCTAGTAAAGACAGAT 513
QY 361 GAGGCAAGATTTAAGGCACTTTTGAAGAAAGAGGCTACACTGTAGTGAATCAAGGT 420
DB 514 GAGGCAAGATTTAAGGCACTTTTGAAGAAAGAGGCTACACTGTAGTGAATCAAGGT 573
QY 421 CAGGTGAAGTATGAGAGCACTTCAGATTCGTTTATTCAGGTCAAGCAAGCTTCTGTT 480

DB 574 CAGAGAAATATGAGAGACCACTCCAGATTCCGTTTATTCAGTCAAGCAAGCTTCTGTT 633
QY 481 GGCACTGAGATATTTTGGGAAAGATCCCAAGATCTGTTTGAAGATGAGCTTTGCCA 540
DB 634 GGCACTGAGATATTTTGGGAAAGATCCCAAGATCTGTTTGAAGATGAGCTTTGCCA 693
QY 541 TTATTTGAGAAAGCTGAGACTATATGAGATCTCGTTTATGATGATCCGCTCACTGCT 600
DB 694 TTATTTGAGAAAGCTGAGACTATATGAGATCTCGTTTATGATGATCCGCTCACTGCT 753
QY 601 CTCACAGAGATTATGAGTTTGTCACTTTTGTACAAAGAACAGACACAGAGGCTGT 650
DB 754 CTCACAGAGATTATGAGTTTGTCACTTTTGTACAAAGAACAGACACAGAGGCTGT 813
QY 661 AAACGTATATATATCATGAAATTCGTTCCGGAGACACATGTGTCTGATCTCATGTT 720
DB 814 AAACGTATATATATCATGAAATTCGTTCCGGAGACACATGTGTCTGATCTCATGTT 873
QY 721 GCCAAGATAGGCTTTTGTGGGCTGATCTTAAGATTAACCAAGAGGAGATTCTT 780
DB 874 GCCAAGATAGGCTTTTGTGGGCTGATCTTAAGATTAACCAAGAGGAGATTCTT 933
QY 781 GAGGAATTTAGCAAGTACAGAGGCTCTCAAGATGTCATTTTATACCAACCACTGAT 840
DB 934 GAGGAATTTAGCAAGTACAGAGGCTCTCAAGATGTCATTTTATACCAACCACTGAT 993
QY 841 GACAGAAAAAAACAGAGGCTTTGCTTTCTTGAATATGAGATCACAAAACAGCTGCC 900
DB 994 GACAGAAAAAAACAGAGGCTTTGCTTTCTTGAATATGAGATCACAAAACAGCTGCC 1053
QY 901 CAGGCAAGAGTGGGCTTAAGATGTTAAAGTCAAAAGTCTGGGAAATGTTGAACTGTT 960
DB 1054 CAGGCAAGAGTGGGCTTAAGATGTTAAAGTCAAAAGTCTGGGAAATGTTGAACTGTT 1113
QY 961 GAGTGGCTGATCTTATGAGATCTGATCTGAAATTATGCAAAAGTAAAGTCTG 1020
DB 1114 GAGTGGCTGATCTTATGAGATCTGATCTGAAATTATGCAAAAGTAAAGTCTG 1173
QY 1021 TTTGTACGCACTTGTCCAAACAGGTACAGAGAAATTTAGAAAGTCAATTAGTCA 1080
DB 1174 TTTGTACGCACTTGTCCAAACAGGTACAGAGAAATTTAGAAAGTCAATTAGTCA 1233
QY 1081 TTTGGGAACTGGAACAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1140
DB 1234 TTTGGGAACTGGAACAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1293
QY 1141 AGAGATGCTGCTCAAGGCTATGAGAAAGATGATGATGATGATGATGATGATGAT 1200
DB 1294 AGAGATGCTGCTCAAGGCTATGAGAAAGATGATGATGATGATGATGATGATGAT 1353
QY 1201 ATTGAATTTGTTTGTAAAGCCACAGATCAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1354 ATTGAATTTGTTTGTAAAGCCACAGATCAGAGAGAGAGAGAGAGAGAGAGAG 1413
QY 1261 CAAGCAGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1320
DB 1414 CAAGCAGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1473
QY 1321 CCCCACAAAGAGGTGAGGCGTGAAGGATGAGAGTGTGATATGATATCTCCAGATTAT 1380
DB 1474 CCCCACAAAGAGGTGAGGCGTGAAGGATGAGAGTGTGATATGATATCTCCAGATTAT 1533
QY 1381 TATGATACGAGATTTATATGATTTATGATTTATGATTTATGATTTATGATTTGA 1440
DB 1534 TATGATACGAGATTTATATGATTTATGATTTATGATTTATGATTTATGATTTGA 1593
QY 1441 TATGAAGATCCATATGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 1500
DB 1594 TATGAAGATCCATATGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 1653
QY 1501 AGAGAGCAAGGAGTGTGCTCCATCCAGAGTCTGAGGCTGCTCCCTCCCTGATGA 1560
DB 1654 AGAGAGCAAGGAGTGTGCTCCATCCAGAGTCTGAGGCTGCTCCCTCCCTGATGA 1713


```

Qy 1081 TTGGGAAACGAGACGAGTAAGAAAGCTAAAGATATGCTTTCATTCATTTTGATGAG 1140
Db 1081 TTGGGAAACGAGACGAGTAAGAAAGCTAAAGATATGCTTTCATTCATTTTGATGAG 1140
Qy 1141 AGAGATGCTGCTCAAGGCTATGAGAAAGTAAGATGTAAGAATTGAGAGGAAAT 1200
Db 1141 CGAGATGCTGCTCAAGGCTATGAGAAAGTAAGATGTAAGAATTGAGAGGAAAT 1200
Qy 1201 ATTGAATGCTTTTGTAGCCACGATCAGAAAGGAAAGAAAGAAAGCTCAGAG 1260
Db 1201 ATTGAATGCTTTTGTAGCCACGATCAGAAAGGAAAGAAAGAAAGCTCAGAG 1260
Qy 1261 CAGACGAAAGATCAATATGATGATGATTAATCACTATGCTCACTCATATGCT 1320
Db 1261 CAGACGAAAGATCAATATGATGATGATTAATCACTATGCTCACTCATATGCT 1320
Qy 1321 CCCCCAAGAGAGTCAAGGCGCTGAGAGTGAAGAGTGGCTATGCTCAGATTA 1380
Db 1321 CCCCCAAGAGAGTCAAGGCGCTGAGAGTGAAGAGTGGCTATGCTCAGATTA 1380
Qy 1381 TATGATAGAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1440
Db 1381 TATGATAGAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1440
Qy 1441 TATGATAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1500
Db 1441 TATGATAGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1500
Qy 1501 AGAGAGCAAGAGGAGTGTCTCCATCCAGAGGCTGTGCTCTCCCGGAGTGA 1560
Db 1501 AGAGAGCAAGAGGAGTGTCTCCATCCAGAGGCTGTGCTCTCCCGGAGTGA 1560
Qy 1561 GCGGTTATCAAGAGAGAGGCGCTGATCAGCAAGAGGCGCTGCTCCCGGAGTGA 1620
Db 1561 GCGGTTATCAAGAGAGAGGCGCTGATCAGCAAGAGGCGCTGCTCCCGGAGTGA 1620
Qy 1621 GGTGCCCAAGAGAGGCGCGGAGGAGAAAGGAGGCTGAGGCGGCTGCTGTTA 1680
Db 1621 GGTGCCCAAGAGAGGCGCGGAGGAGAAAGGAGGCTGAGGCGGCTGCTGTTA 1680
Qy 1681 CAA 1683
Db 1681 CAA 1683

```

RESULT 4
AF155568 2221 bp mRNA linear PRI 10-JUN-1999
DEFINITION Homo sapiens NSAP1 protein (NSAP1) mRNA, complete cds.
ACCESSION AF155568
VERSION AF155568.1 GI:5031511
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2221)
AUTHORS Harris,C.E., Boden,R.A. and Astell,C.R.
TITLE A novel heterogeneous nuclear ribonucleoprotein-1-like protein
interacts with NS1 of the minute virus of mice
JOURNAL J. Virol. 73 (1), 72-80 (1999)
MEDLINE 99102562
PUBMED 9847309
REFERENCE 2 (bases 1 to 2221)
AUTHORS Harris,C.E., Boden,R.A. and Astell,C.R.
TITLE Direct Submision
JOURNAL Submitted (01-JUN-1999) Biochemistry and Molecular Biology, UBC,
2146 Health Sciences Mall, Vancouver, BC V6T 1Z3, Canada
FEATURES
source
1..2221
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/note="cell line"
1..2221
/gene="NSAP1"
205..1893
/feature="NSAP1"
/note="similar to Homo sapiens Hn RNP R; probable RNA
binding protein; interacts with NS1 of the minute virus of
mice"
/codon_start=1
/product="NSAP1 protein"
/protein_id="A038198.1"
/db_xref="GI:5031512"
/translation="MATEHWNGNGTBBPMDTTSVAVHSNPOTLLDAGLPKVAEKLD
RIVVAGLVASHSDLRATBALKEPMDALVLIQPKSDLSHONKSAFLGCVKTY
RORRQGTNRVADSSKPPBAKIKALLETGTTLVDVTRQKRGFPSPVSYSQDPSV
GTEIFVGKIPDLFEDLVPLFEKAGPIWDLRLMDPLGLNRGAFLFCTKEA0E
AVKLVNNEHRSKGIIVCISVANNRLPVGSIKSKTE0LLEBSF0VTEGLDVLV
HOPDCKNRGFCLEEDHKTAAQARRLMSGKVKVNGVNGVEMADIEDPDEEM
AKVKVLFVRLANTVTEELIEKAP0F0KLRVVKLKYAFIHPERDGAIVAMEEM
KQDLGENIETVFAKPPDQKREKRA0QA0KNQYDYIYTPPHMPPTRGRGCG
RGVGYPPDYGYEDYDYDYDYDHNHYGVEDPYGYEDYOVGARGRGARGAAP
SRKGAAPPRGAGYS0RGPSASGVARGAGA0Q0R0G0G0K0G0V0E0P0D0L0Q0"
BASE COUNT 699 a 402 c 561 g 559 t
ORIGIN
Query Match 93.5%; Score 1574; DB 9; Length 2221;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 60; Indels 3; Gaps 1;
Qy- 1 ATGGCTACAGAACATGTATAGTAAGAAATGTACTGAAGAGCCCATGGAATCTACGCA 60
Db 205 ATGGCTACAGAACATGTATAGTAAGAAATGTACTGAAGAGCCCATGGAATCTACGCA 264
Qy 61 GTTATCATTCAGAAAATTTTCAAGACATTCCTGATGCTGTGTTTACACAGAAAGTTGCT 120
Db 265 GTTATCATTCAGAAAATTTTCAAGACATTCCTGATGCTGTGTTTACACAGAAAGTTGCT 324
Qy 121 GAAAACTGAGAAATTTTACGTTGAGAGGCTAGTTCACATAGTGAATTAATGAAAGA 180
Db 325 GAAAACTGAGAAATTTTACGTTGAGAGGCTAGTTCACATAGTGAATTAATGAAAGA 384
Qy 181 GCTATCAAGCTTTTAAAGATTCATGAAAGACGGGCAATGGGAGTGGCTTCAACAGTTT 240
Db 385 GCTATCAAGCTTTTAAAGATTCATGAAAGACGGGCAATGGGAGTGGCTTCAACAGTTT 444
Qy 241 AAAGACAGTATCTCTCATGTTTCAAGAAAGAGTGCCTTTTATGTGAGTCAATGAG 300
Db 445 AAAGACAGTATCTCTCATGTTTCAAGAAAGAGTGCCTTTTATGTGAGTCAATGAG 504
Qy 301 ACTTACAGCAGAGAGAAACAGAGGACCAAGTACAGTCTATGTAAGACCAAT 360
Db 505 ACTTACAGCAGAGAGAAACAGAGGACCAAGTACAGTCTATGTAAGACCAAT 564
Qy 361 GAGGCAAGATTAAGGCACTTTTGGAAAGACAGGCTACACCTGATGATCAAGT 420
Db 565 GAGGCAAGATTAAGGCACTTTTGGAAAGACAGGCTACACCTGATGATCAAGT 624
Qy 421 CAGAGAGATGAGAGACCACTCCAGATTCGTTTATTCAGTCAAGAGCTTCTGTT 480
Db 625 CAGAGAGATGAGAGACCACTCCAGATTCGTTTATTCAGTCAAGAGCTTCTGTT 684
Qy 481 GGCACCTGAGATTTTGTGGGAAAGATCCCAAGAGATCTGTTTGAAGATGAGCTTGTCCA 540
Db 685 GGCACCTGAGATTTTGTGGGAAAGATCCCAAGAGATCTGTTTGAAGATGAGCTTGTCCA 744
Qy 541 TATTTGAGAAAGCGGACCTATAGGAGATCTTGTTAATGATGATGATGATGATGAT 600
Db 745 TATTTGAGAAAGCGGACCTATAGGAGATCTTGTTAATGATGATGATGATGATGAT 804
Qy 601 CTCACAGAGGTTATGCTTGTCACTTTTGTACAAAAGAGACAGACAAAGGCTGTT 660
Db 805 CTCACAGAGGTTATGCTTGTCACTTTTGTACAAAAGAGACAGACAAAGGCTGTT 864

```

QY 661 AAATGTAATATATCATGAAATTCGTCGGGAAGCAATTGGTGTCTGATCTCACTT 720
 DB 865 AAATGTAATATATCATGAAATTCGTCGGGAAGCAATTGGTGTCTGATCTCACTT 924
 QY 721 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAGAGACAGATTCCT 780
 DB 925 GCCAACAATAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAGAGACAGATTCCT 984
 QY 781 GAGGAATTTAGCAAGAGACAGAGAGGCTCAAGATGCTATTTATACCAACCTGAT 840
 DB 985 GAGGAATTTAGCAAGAGAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAGAGAC 1044
 QY 841 GACAAGAAAAAACAAGAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAGAGAC 900
 DB 1045 GACAAGAAAAAACAAGAGGCTTTTGTGGGCTCGATTCCTAAGATAAACCAGAGAC 1104
 QY 901 CAGGCAAGACCTAGGCTTAATAGTGTCTAAGTCAAAAGCTCTGGGAAATGTTGAACTGTT 960
 DB 1105 CAGGCAAGACCTAGGCTTAATAGTGTCTAAGTCAAAAGCTCTGGGAAATGTTGAACTGTT 1164
 QY 961 GAGTGGGCTGATCTCTATTGAAGATCTGATCTGAAAGTTATGCGAAAGGTAAGTCTG 1020
 DB 1165 GAGTGGGCTGATCTCTATTGAAGATCTGATCTGAAAGTTATGCGAAAGGTAAGTCTG 1224
 QY 1021 TTGTAGCGAACCTTGCCAAACAGGCTAACAGAAATTTAGAAAGTCAATTAGTCTAG 1080
 DB 1225 TTGTAGCGAACCTTGCCAAACAGGCTAACAGAAATTTAGAAAGTCAATTAGTCTAG 1284
 QY 1081 TTGTGGAACTGGAACAGAGTAAAGAGTAAAGTTATGCTTTCATTCATTTGATGAG 1140
 DB 1285 TTGTGGAACTGGAACAGAGTAAAGAGTAAAGTTATGCTTTCATTCATTTGATGAG 1344
 QY 1141 AGAGATGCTGCTGTCAGAGGCTATGAGAAATGATGCTTAAGACTTGAGAGGAGAAAT 1200
 DB 1345 AGAGATGCTGCTGTCAGAGGCTATGAGAAATGATGCTTAAGACTTGAGAGGAGAAAT 1464
 QY 1201 ATTGAATTTGTTTTGCTTGAACCAACAGATCAGAAAGAGAAAGAAAGTCTCAGAG 1260
 DB 1405 ATTGAATTTGTTTTGCTTGAACCAACAGATCAGAAAGAGAAAGAAAGTCTCAGAG 1464
 QY 1261 CAAACAGAAAGAAATCAAAATGATGATGATTAATGCTTCACTATATGCTCT 1320
 DB 1465 CAAACAGAAAGAAATCAAAATGATGATGATTAATGCTTCACTATATGCTCT 1524
 QY 1321 CCCCACAAAGAGGTCGAGGCGCTGAGAGTGAAGGTCGATGATCTCTCAGATTAAT 1380
 DB 1525 CCCCACAAAGAGGTCGAGGCGCTGAGAGTGAAGGTCGATGATCTCTCAGATTAAT 1584
 QY 1381 TATGATTCAGAAATTTATGATTTATGATGATTAATGATTAATGATTAATGATTAAT 1440
 DB 1585 TATGATTCAGAAATTTATGATTTATGATGATTTATGATTAATGATTAATGATTAAT 1644
 QY 1441 TATGAAGATTCATTAATGATTTATGATGATTTATGATTAATGATTAATGATTAAT 1500
 DB 1645 TATGAAGATTCATTAATGATTTATGATGATTTATGATTAATGATTAATGATTAAT 1704
 QY 1501 AGAGAGAGAGAGGCTGCTGCTCAACAGAGGTCGAGGCTCTCTCTCTCTCTCTCTCT 1560
 DB 1705 AGAGAGAGAGAGGCTGCTGCTCAACAGAGGTCGAGGCTCTCTCTCTCTCTCTCTCT 1764
 QY 1561 GCCGATTTATTCAGAGAGAGAGGCTCTGATCAGCAAGAGGCTCTCTCTCTCTCTCTCT 1620
 DB 1765 GCCGATTTATTCAGAGAGAGAGGCTCTGATCAGCAAGAGGCTCTCTCTCTCTCTCTCT 1824
 QY 1621 GGTGCCCCAACAAAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1677
 DB 1825 GGTGCCCCAACAAAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1884
 QY 1678 TTACAA 1683
 DB 1885 TTACAA 1890

RESULT 5.
 AY034481 1872 bp mRNA linear PRI 03-OCT-2001
 LOCUS AY034481
 DEFINITION Homo sapiens hnRNP Q3 mRNA, complete cds.
 ACCESSION AY034481
 VERSION AY034481.1 GI:15809585
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1872)
 Mounelacos, Z., Abel, L., Yong, J., Kataoka, N. and Dreyfuss, G.
 SMN interacts with a novel family of hnRNP and spliceosomal
 proteins
 JOURNAL EMBO J. 20 (19), 5443-5452 (2001)
 MEDLINE 21458434
 PUBMED 11574476
 REFERENCE
 2 (bases 1 to 1872)
 Mounelacos, Z.
 Direct Submision
 Submitted (11-MAY-2001) HMI/ Biochemistry, University of
 Pennsylvania, 330 CRB, 415 Curie Blvd., Philadelphia, PA 19104, USA
 FEATURES
 source
 1..1872
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..1872
 /codon_start=1
 /product="hnRNP Q3"
 /protein_id="AAK59703.1"
 /db_xref="GI:15809586"
 /translation="MATEHVNGNGTBEPMDDTSAAVISHENFOTLLDAGLPQVAREKLD
 ELYVAGLVASDLDERRALEAKFENEDGALVLOPKDSLSHVONKSAFLGVWKTY
 ROREKQGTVAQDSKGPDEAKIKALLERTGYLDVTTGQRKYGGPPDSVSGQPSV
 GTEIFVYKPIRDLFEDLVPFEKAGPIWDLMLMDPLGLRGAFTVFCREKAOB
 AVKLVNHEIRBSGKHIGVCI SVANNRLPGSI PKSTKBOLEBSKTEGLTDVILY
 HPPDDKKRSFCFLIEDHKTAQARRRLNSGKRYKWNQVETVEMADPIEDPBEVM
 ARVKLVFNVNLTAVTEELERAFQFGLGVKLVKLDYAFIHFPERDAVAKBEMN
 GDLBGENEILVFAKPPDOKRKERAKROAONWYDDVYVYVPHMPPVTEBRSG
 RGVGYPPDYEDYDYDYDYDYNHRYGVDYDYEDYFVGARGGGARGARALP
 SRGRGAAPRGARAGYSORGGPSGARVGARAGAGAOORGVARGARGGGARGK
 AGVNDPDSKRGQTNNQWGSOPINQOPLGGSDHSGNIVKSENEFYDPTDTCQWQK"
 BASE COUNT 590 a 323 c 489 g 470 t
 ORIGIN
 Query Match 92.2%; Score 1552.2; DB 9; Length 1872;
 Best Local Similarity 95.6%; Pred. No. 0;
 Matches 1596; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 1 ATGGCTACAGAACATGTTAATGGAATGCTGTAAGAGCCCATGATTAATCTTACACA 60
 DB 1 ATGGCTACAGAACATGTTAATGGAATGCTGTAAGAGCCCATGATTAATCTTACACA 60
 QY 61 GTTATTCATTCAGAAAATTTTTCAGACATGCTGATGCTGCTTACACAGAAAGTTGCT 120
 DB 61 GTTATTCATTCAGAAAATTTTTCAGACATGCTGATGCTGCTTACACAGAAAGTTGCT 120
 QY 121 GAAAACTAGATAAATTTAGCTTGAAGGCTAGTTGACATATGATTTAGATGAAGA 180
 DB 121 GAAAACTAGATAAATTTAGCTTGAAGGCTAGTTGACATATGATTTAGATGAAGA 180
 QY 181 GCTATCAGACCTTTAAAAGATTCAATGAAGACGGGCTATGGCAGTCTTCAACAGTTT 240
 DB 181 GCTATCAGACCTTTAAAAGATTCAATGAAGACGGGCTATGGCAGTCTTCAACAGTTT 240
 QY 241 AAGAGACGATATCTCTCTCATGTTTCAGAACAAAGAGCCCTTTTATGAGAGCAAGAAG 300
 DB 241 AAGAGACGATATCTCTCTCATGTTTCAGAACAAAGAGCCCTTTTATGAGAGCAAGAAG 300
 QY 301 ACTTACAGCAGAGAGAAAACAGGGGACCAAGATGACAGACTCTTGAAGAGACAGAT 360
 DB 301 ACTTACAGCAGAGAGAAAACAGGGGACCAAGATGACAGACTCTTGAAGAGACAGAT 360

OY	1	ATGGCTACAGAAACATGTTAATGAGAAATGTAACAGAGGCCATGGAATCTACTCTCACAG	60
Db	526	ATGGCTACAGAAACATGTTAATGAGAAATGTAACAGAGGCCATGGAATCTACTCTCACAG	585
OY	61	GTTATCCATTCAGAAAAATTTTTCAGACATTCCTGATGATGCTGGTTTACACAGAAAGTTGCT	120
Db	586	GTTATCCATTCAGAAAAATTTTTCAGACATTCCTGATGATGCTGGTTTACACAGAAAGTTGCT	645
OY	121	GAAAAACTAGATGAAATTTTACGTTGACGGGCTTACTGACATATGTAATTTAGTAAGAA	180
Db	646	GAAAAACTAGATGAAATTTTACGTTGACGGGCTTACTGACATATGTAATTTAGTAAGAA	705
OY	181	GCTATCGAAGCTTTTAAAAAGATTCAATGAAAGACGGGCAATTGGCAAGTCTTCAAGATT	240
Db	706	GCTATCGAAGCTTTTAAAAAGATTCAATGAAAGACGGGCAATTGGCAAGTCTTCAAGATT	765
OY	241	AAAGACAGTGAATCTCTCATGTTTCAGAACAAAGTGCCCTTTTATATGTAAGTCAATGAG	300
Db	766	AAAGACAGTGAATCTCTCATGTTTCAGAACAAAGTGCCCTTTTATATGTAAGTCAATGAG	825
OY	301	ACTTACAGGCGAAGAAAAACAGGGGACAAAGCTATGACGACTTATATGTAAGACCAAGAT	360
Db	826	ACTTACAGGCGAAGAAAAACAGGGGACAAAGCTATGACGACTTATATGTAAGACCAAGAT	885
OY	361	GAGGCAAAAGATTAAAGCACTTTTGGAAAGAACAGGCTCAACACTGATGTCATACAGAT	420
Db	886	GAGGCAAAAGATTAAAGCACTTTTGGAAAGAACAGGCTCAACACTGATGTCATACAGAT	945
OY	421	CAGAGAGATATGAGAGACCACTCCAGATTCGGTTATTCAGGTACAGCAAGCTCTGT	480
Db	946	CAGAGAGATATGAGAGACCACTCCAGATTCGGTTATTCAGGTACAGCAAGCTCTGT	1005
OY	481	GGCACTGAGATATTTTGGGGAGAGATCCCAAGATCTGTTGAGATGAGCTTGTTCOA	540
Db	1006	GGCACTGAGATATTTTGGGGAGAGATCCCAAGATCTGTTGAGATGAGCTTGTTCOA	1065
OY	541	TTATTTGAGAAAGCTGAGCCATATGAGGATCTTGTTATATATGATATCCGCTCACTGAT	600
Db	1066	TTATTTGAGAAAGCTGAGCCATATGAGGATCTTGTTATATATGATATCCGCTCACTGAT	1125
OY	601	CTCAACAGAGGTTATGCGTTTGTCACTTTTGTACAAAAGAGCAGACAAAGAGCTGTT	660
Db	1126	CTCAATGAGGTTATGCGTTTGTCACTTTTGTACAAAAGAGCAGCTCAGAGAGCTGTT	1185
OY	661	AAACTGTATTAATTAATCATGAAATTCGTTCCGAGAGCAATTGCTGTCACTCAATT	720
Db	1186	AAACTGTATTAATTAATCATGAAATTCGTTCCGAGAGCAATTGCTGTCACTCAATT	1245
OY	721	GCCAAACAATAGGCTTTTGTGGGCTCGAATCCATAGATGTAACCAAGAGAGATTTT	780
Db	1246	GCCAAACAATAGGCTTTTGTGGGCTCGAATCCATAGATGTAACCAAGAGAGATTTT	1305
OY	781	GAGNAATTTAGCAAGTGAAGAGAGGGCTCAAGATGTCAATTTTATACCAACCACTGAT	840
Db	1306	GAGNAATTTAGCAAGTGAAGAGAGGGCTTATCAGAGCTCAATTTTATACCAACCACTGAT	1365
OY	841	GACAAAGAAAAAACAAGAGCTTTTGTCTTCTTGAATATGAAATCACAAAACAGTGC	900
Db	1366	GACAAAGAAAAAACAAGAGCTTTTGTCTTCTTGAATATGAAATCACAAAACAGTGC	1425
OY	901	CAGGCAAGACCTTAGGCTTATGATGATGCTTAAAGTCTGGGAGAAATGTGGAACTGTT	960
Db	1426	CAGGCAAGGCTTAGGCTTATGATGATGCTTAAAGTCTGGGAGAAATGTGGAACTGTT	1485
OY	961	GAGTGGGCTGATCTCATTTGAAGATCTCGATCTGAAAGTTATGGCAAGGTAAAGTGTG	1020
Db	1486	GAGTGGGCTGATCTCATTTGAAGATCTCGATCTGAAAGTTATGGCAAGGTAAAGTGTG	1545
OY	1021	TTTGTACGCAACTTGCACAACGGTAAACAGAAATTTTGAAGAAAGCTATTAGTCAG	1080
Db	1546	TTTGTACGCAACTTGCACAATCTGTAAACAGAAAGATTTTGAAGAAAGCTATTAGTCAG	1605
OY	1081	TTTGGGAAACTGGAACGATTAAGAGCTTAAAGATTAATGCTTCAATTCATTTGATGATG	1140

Db	1606	TTTGGAAATCTGAAACGAGTGAAGAAAGTTAAAGATTATGCTTCAATTTTGATGAG	16655
Oy	1141	AGAGATGTCCTGTCAAAGCTATGAAAGAAATGAATGTAAAGCTTGAGGGAGAAAT	12000
Db	1666	CGAGATGTCCTGTCAAAGCTATGAAAGAAATGAATGTGCAAGACTTGAGAGGAGAAAT	17255
Oy	1201	ATTGAATGTGTTTTTGCTAAAGCCACGATCGAAGAGGAAAGAAAGAAAGCTCAGAG	12606
Db	1726	ATTGAATGTGTTTTTGCTAAAGCCACGATCGAAGAGGAAAGAAAGAAAGCTCAGAG	17855
Oy	1261	CAAGCAGCAAGATCAATGATGATGATCTACTATATGTCACCTCATATGCT	13200
Db	1786	CAAGCAGCAAGAAATCAATGATGATGATCTACTATATGTCACCTCATATGCTC	18455
Oy	1321	CCCCCAAGAAGTGAAGGGCGTGAAGGTGAAGGTGCTATGATATCTTCAGATTAT	13800
Db	1846	CTTCACCAAGAGGTGAAGGGCGTGAAGGTGAAGGTGCTATGATATCTTCAGATTAT	19055
Oy	1381	TATGATACGAAGATTATATGATTTATATGTTATGATTACATATCATTCGTGTGGA	14400
Db	1906	TATGATACGAAGATTATATGATTTATATGTTATGATTATACATATCATTCGTGTGGA	19655
Oy	1441	TATGAATCATATATGTTATGAAATTTTCAAGTTGAGCTAGAGAAAGGGTGT	15000
Db	1966	TATGAATCATATATGTTATGAAAGTTTCAAGTTGAGCTAGAGAAAGGGTGT	20255
Oy	1501	AGAGAGCAAGGGGTGTGTCATCCAAAGTCGTGGGCGTCTCCCTCCGTGTGA	15606
Db	2026	AGAGAGCAAGGGGTGTGTCATCCAAAGTCGTGGGCGTCTCTCCCGCGGTGA	20855
Oy	1561	GCCGGTATTTCAGAGAGAGGCGCTTGATGCAAGAGGCGTTCGCGTCCGAGAGA	16200
Db	2086	GCCGGTATTTCAGAGAGAGGCGCTTGATGCAAGAGGCGTTCGAGGTCCGAGAGA	21455
Oy	1621	GCTGCCCAACAAAGAGGCGCGGGGGGAAAGGGGTCTAGAGCCGCTC	1669
Db	2146	GCTGCCCAACAAAGAGGCGCGGGGGTATCGTGTGCGAGGGGTGGCC	2194

RESULT 7	2079 bp	DNA	linear	PAT 17-DEC-2001
AR177851				
LOCUS				
DEFINITION	Sequence 2 from patent US 6313266.			
ACCESSION	AR177851			
VERSION	AR177851.1			
KEYWORDS	GI:17920206			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 2079)			
TITLE	Bandman,O., Yue,H., Corley,N.C. and Shah,P.			
JOURNAL	Human nucleolin-like protein			
FEATURES	Patent: US 6313266-A 2 06-NOV-2001;			
source	Location/Qualifiers			
	1..2079			
	/organism="unknown"			
BASE COUNT	639 a 377 c 535 g 528 t			
ORIGIN				
Query Match	90.3%; Score 1519; DB 6; Length 2079;			
Best Local Similarity	96.0%; Pred. No. 0;			
Matches 1558; Conservative	0; Mismatches 65; Indels 0; Gaps 0.			
OY	1 ATGGCTACAGAACATGTTAATGGAATGGTACTGAAAGCCCATGAGTACTTACTTCAGCA 60			
Db	203 ATGGCTACAGAACATGTTAATGGAATGGTACTGAAAGCCCATGAGTACTTACTTCAGCA 262			
OY	61 GTTATCATTCAGAAAATTTTCAGACATTGGCTGATGCGGTTTACACAGAAAGTTGCT 120			
Db	263 GTTATCATTCAGAAAATTTTCAGACATTGGCTGATGCGGTTTACACAGAAAGTTGCT 322			
OY	121 GAAAACTAGATGAATTTACGTTGCAGGGCGTAGTTGCACATAGTAGATTAGATGAAAGA 180			

Db 323 GAAAACTAGATGAAATTTAGCTTGGAGGCTAGTTGCATCATAGTATTAAGTAAAGA 382
 Qy 181 GCTATGAGACCTTTAAAGAGTCAATGAAGACGGCCGATTTGGAGAGCTTCAACAGTTT 240
 Db 383 GCTATGAGACCTTTAAAGAGTCAATGAAGACGGTGCATTTGGAGAGCTTCAACAGTTT 442
 Qy 241 AAGACAGTATCTCTCATGTTTCAAGAACAAAGTCCCTTTTATATGAGAGTCAATGAG 300
 Db 443 AAGACAGTATCTCTCATGTTTCAAGAACAAAGTCCCTTTTATATGAGAGTCAATGAG 502
 Qy 301 ACTTACAGCAG 360
 Db 503 ACTTACAGCAG 562
 Qy 361 GAGGCAAGATTAAAGGACCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Db 563 GAGGCAAGATTAAAGGACCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 622
 Qy 421 CAGAGAGAGATGAG 480
 Db 623 CAGAGAGAGATGAG 682
 Qy 481 GGCAGTGAATATTTTGGAG 540
 Db 683 GGCAGTGAATATTTTGGAG 742
 Qy 541 TATTTGAG 600
 Db 743 TATTTGAG 802
 Qy 601 CTCACAGAGAGATGAG 660
 Db 803 CTCACAGAGAGATGAG 862
 Qy 661 AAACGTAT 720
 Db 863 AAACGTAT 922
 Qy 721 GCCAACAT 780
 Db 923 GCCAACAT 982
 Qy 781 GAGGAATTTAGCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 Db 983 GAGGAATTTAGCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1042
 Qy 841 GACCAAG 900
 Db 1043 GACCAAG 1102
 Qy 901 CAGGCAAG 960
 Db 1103 CAGGCAAG 1162
 Qy 961 GAGTGGAGTATCTCTTATGAAGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGAT 1020
 Db 1163 GAGTGGAGTATCTCTTATGAAGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGAT 1222
 Qy 1021 TTTGTACGCAACCTTGCACAAACGCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 Db 1223 TTTGTACGCAACCTTGCACAAACGCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1282
 Qy 1081 TTTGGAGAACTGGAACGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 Db 1283 TTTGGAGAACTGGAACGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1342
 Qy 1141 AGAGATGAGTCTCTCAAGAGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 Db 1343 CGAGATGAGTCTCTCAAGAGTATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1402
 Qy 1201 ATTGAATTTGTTTGTAAAGCAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

Db 1403 ATTGAATTTGTTTGTCCAGAGCAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1462
 Qy 1261 CAG 1320
 Db 1463 CAG 1522
 Qy 1321 CCCCAG 1380
 Db 1523 CCCCAG 1582
 Qy 1381 TATGATGAG 1440
 Db 1583 TATGATGAG 1642
 Qy 1441 TATGATGAG 1500
 Db 1643 TATGATGAG 1702
 Qy 1501 AG 1560
 Db 1703 AG 1762
 Qy 1561 GCCGTTATTTCAAG 1620
 Db 1763 GCCGTTATTTCAAG 1822
 Qy 1621 GGT 1623
 Db 1823 GGT 1825

RESULT 8
 AF408434 2232 bp mRNA linear ROD 10-OCT-2001
 AF408434
 LOCUS Mus musculus RRM RNA binding protein NSAP1 (Nsap1) mRNA, partial
 DEFINITION
 ACCESSION AF408434
 VERSION AF408434.1 GI:15986734
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerogomphini; Muridae; Murinae; Mus.
 1 (bases 1 to 2232)
 Zhou, M. and Raschke, W.C.
 TITLE
 Mus musculus RRM RNA binding protein, NSAP1
 JOURNAL
 Unpublished
 2 (bases 1 to 2232)
 Zhou, M. and Raschke, W.C.
 REFERENCE
 Direct Submission
 Submitted (13-AUG-2001) Immunology, Sidney Kimmel Cancer Center,
 10835 Altman Row, San Diego, CA 92121, USA
 JOURNAL
 Location/Qualifiers
 1..2232
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 <1..2232
 /gene="Nsap1"
 <1..1679
 /gene="Nsap1"
 /codon_start=3
 /product="RRM RNA binding protein NSAP1"
 /protein_id="AA11726.1"
 /db_xref="GI:15986734"
 translation="KKEPNDGALVLAOFKSDLSHYONKSAFLGWNKYRORPKOG
 TKVDSGSGPBEAKIKLLERTGTTDITDQGRKGPFPDPSVYSGGQPSVGTETLFG
 KIPRLDEDELVPLFEKAGPTMDLMDPLTGLNRGTAFTCTKEAQAQAVKLYNN
 HIRSGKHIIVGICISVANNRLVVGSIPIKSTKEQLLEBSKYTBGLTDVILVHOPDKK
 NKRGCFLVEDHDKTAQARRRLMSGKTVGANTGVEMADPIEDPDEVAKYVLF
 VRNLANTYEEIIEKSFQFGLSRVKKLKOYAFIHPERDGAIVAMEEMKQKLEGE
 NIEIVFAKPPQKRRKROAKKNQYDIDYVYGGPHMPPTRGRGGRGGAGVGP
 PDYGVENYVYVGYDYNRYGVEDPYVYGEDPYONKARGGSGARGAASRSRGAA
 PPRGRAGISQNGPESANRVARGAQOQNGRGRGRGRGAGVGGKRYADGNP

DTKRRNNQNMGSOPIAOQPLGGSDHSNGYKSENOBFYDTPGQWKK"
 BASE COUNT 689 a 387 c 551 g 605 t

Query Match 86.6%; Score 1456.8; DB 10; Length 2232;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1461; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 194 TAAAGAGTTCAATGAGAGCGGCGCATTTGGAGTCTTCAACAGTTTAAAGACAGTATC 253
 DB 1 TAAAGAGTTCAATGAGAGCGGCGCATTTGGAGTCTTCAACAGTTTAAAGACAGTATC 60
 QY 254 TCTCTCATGTTCAAGACAAAGTGCCTTTTATGAGTATGAGTAAAGATCTACAGGCGA 313
 DB 61 TCTCTCATGTTCAAGACAAAGTGCCTTTTATGAGTATGAGTAAAGATCTACAGGCGA 120
 QY 314 GAGAAAAACAGGGGACCAAGTAGACAGCTAGTAAAGACAGATGAGCAAGATTA 373
 DB 121 GAGAAAAACAGGGGACCAAGTAGACAGCTAGTAAAGACAGATGAGCAAGATTA 180
 QY 374 AGGCACTTTTGAAGAAAGACAGGCTACAGCTTATGATGATCTACAGTCAAGAGAAATG 433
 DB 181 AGGCACTTTTGAAGAAAGACAGGCTACAGCTTATGATGATCTACAGTCAAGAGAAATG 240
 QY 434 GAGAGCACTTCCAGATTCGCTTATTCAGGTCAGACGCTTCTGTTGGCATCTAGATAT 493
 DB 241 GAGAGCACTTCCAGATTCGCTTATTCAGGTCAGACGCTTCTGTTGGCATCTAGATAT 300
 QY 494 TTGTGGGAGAGATCCCAAGAGATCTGTTGAGATGAGATGATGATGATGATGATGATG 553
 DB 301 TTGTGGGAGAGATCCCAAGAGATCTGTTGAGATGAGATGATGATGATGATGATGATG 360
 QY 554 CTGAGACCTATATGAGATCTTGTGTTAATGATGATGATGATGATGATGATGATGATG 613
 DB 361 CTGAGACCTATATGAGATCTTGTGTTAATGATGATGATGATGATGATGATGATGATG 420
 QY 614 ATGGCTTGTCACTTTTGTACAAAAGAGACAGCAAGAGGCTGTAACTGTATATA 673
 DB 421 ATGGCTTGTCACTTTTGTACAAAAGAGACAGCAAGAGGCTGTAACTGTATATA 480
 QY 674 ATCATGAATTCGTTCCGGGAGACATATGATGATGATGATGATGATGATGATGATG 733
 DB 481 ATCATGAATTCGTTCCGGGAGACATATGATGATGATGATGATGATGATGATGATG 540
 QY 734 TTTTGTGGGCTCCATTTCTTAAGATTAACCAAGACAGATTTCTTAAGATTTAGCA 793
 DB 541 TTTTGTGGGCTCCATTTCTTAAGATTAACCAAGACAGATTTCTTAAGATTTAGCA 600
 QY 794 AAGTGAAGAGGCTCTCAAGATGCTATTTATACCAACCTGATGACAAAGAAAAA 853
 DB 601 AAGTGAAGAGGCTCTCAAGATGCTATTTATACCAACCTGATGACAAAGAAAAA 660
 QY 854 ACAGAGGCTTTTGTCTTGTGAATATGAGATCACAAAAACAGCTGCCAGCAAGAGTA 913
 DB 661 ACAGAGGCTTTTGTCTTGTGAATATGAGATCACAAAAACAGCTGCCAGCAAGAGTA 720
 QY 914 GGTCAATGAAGTGTAAATCAAGCTGCGGAGAAATGTTGAACTTTAGAGTGGCTATC 973
 DB 721 GGTCAATGAAGTGTAAATCAAGCTGCGGAGAAATGTTGAACTTTAGAGTGGCTATC 780
 QY 974 CTATTGAAGATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1033
 DB 781 CTATTGAAGATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 1034 TTGCCAACAAGGTAACAGAAATTTTGAAGAACTTTAGTCAAGTTGGGAAACTGG 1093
 DB 841 TTGCCAACAAGGTAACAGAAATTTTGAAGAACTTTAGTCAAGTTGGGAAACTGG 900
 QY 1094 AACGAGTAAAGAGCTAAAGATTTATGCTTTTCACTTTTGAAGAGAGATGATGATG 1153
 DB 901 AACGAGTAAAGAGCTAAAGATTTATGCTTTTCACTTTTGAAGAGAGATGATGATG 960
 QY 1154 TCAAGGCTATGGAAGAAATGATGTAAGATCTGAGAGGAGAAATATGTAATTTGTTT 1213

DB 961 TCAAGGCTATGGAAGAAATGATGTAAGATCTGAGAGGAGAAATTTGTAATTTGTTT 1020
 QY 1214 TTGCTAAGCCACCAATCCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGAA 1273
 DB 1021 TTGCTAAGCCACCAATCCAGAGAGAGAAAGAAAGAAAGCTCAGAGGCAAGCAAGAA 1080
 QY 1274 ATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1333
 DB 1081 ATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1334 GTCAAGGCGCTGAGAGTATGAGTGTATGATGATGATGATGATGATGATGATGATG 1393
 DB 1141 GTCAAGGCGCTGAGAGTATGAGTGTATGATGATGATGATGATGATGATGATGATG 1200
 QY 1394 ATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1453
 DB 1201 ATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 QY 1454 ACTATGCTATGAAAGTTTCAAGTTGAGCTAGAGAGGAGGCTGATGAGAGCAAGG 1513
 DB 1261 ACTATGCTATGAAAGTTTCAAGTTGAGCTAGAGAGGAGGCTGATGAGAGCAAGG 1320
 QY 1514 GTGCTGCTCATCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1573
 DB 1321 GTGCTGCTCATCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1574 AGAGAGAGGCGCTGATCAGCAAGAGGCTGCGGTGCGAGAGAGAGTCCCAACAC 1633
 DB 1381 AGAGAGAGGCGCTGATCAGCAAGAGGCTGCGGTGCGAGAGAGAGTCCCAACAC 1440
 QY 1634 AAGAGGCGCGGGGAGAAAGGGGTGAGAGGCGGCTC 1669
 DB 1441 AAGAGGCGCGGGGATCGTGTGCAAGGGGTGCTC 1476

RESULT 9
 HSJ75C13
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP4-775C13 on chromosome 20p12.1-13.
 Contains an hnmnp R (heterogenous nuclear riboprotein R, RRM rna
 binding protein GRY-RBP) pseudogene, ESTs, STS and GSS, complete
 sequence.

ACCESSION
 AL109618
 VERSION
 AL109618.1 GI:5668665
 KEYWORDS
 HTG; GRY-RBP; hnmnp; RRM.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 103819)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Direct Submision
 Submitted (02-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerquest@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30)
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em; EMBL; SW;
 SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormp This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20> RP4-775C13 is from the library RPCT-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pCVPAC2

This sequence is the entire insert of clone RP4-775C13.

Location/Qualifiers

source

```
1. .103819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="p12.1-13"
/clone="RP4-775C13"
/clone_lib="RPCT-4"
1. .236
/note="AluY repeat: matches 1. .236 of consensus"
repeat_region
293. .562
/note="L1M4c repeat: matches 1844. .1624 of consensus"
repeat_region
819. .1471
/note="L1M3c repeat: matches 969. .1634 of consensus"
repeat_region
2177. .2588
/note="L1M3c repeat: matches 14. .431 of consensus"
repeat_region
2593. .2652
/note="MUTIF repeat: matches 474. .533 of consensus"
repeat_region
2701. .3176
/note="L238 copies 2 mer tt 54% conserved"
repeat_region
2814. .2978
/note="L15 copies 11 mer 66% conserved"
repeat_region
2996. .3196
/note="L67 copies 3 mer tcc 56% conserved"
repeat_region
3031. .3184
/note="L14 copies 11 mer 63% conserved"
repeat_region
3296. .4082
/note="L1M4 repeat: matches 3350. .4128 of consensus"
repeat_region
4197. .4315
/note="L1M4 repeat: matches 3021. .3136 of consensus"
repeat_region
4316. .4738
/note="MSTA repeat: matches 1. .426 of consensus"
repeat_region
4739. .5118
/note="L1M4 repeat: matches 2626. .3021 of consensus"
repeat_region
5146. .5288
/note="match: GSS: Em:B90786"
misc_feature
5345. .6619
/note="L1PA16 repeat: matches 3863. .5135 of consensus"
repeat_region
6882. .7139
/note="L129 copies 2 mer gg 58% conserved"
repeat_region
7351. .8891
/gene="dJ775C13.1"
751. .8991
/gene="dJ775C13.1"
CDS
/gene="dJ775C13.1"
/note="dJ775C13.1 (hnRNP R (heterogenous nuclear riboprotein R, RRM rna binding protein GRX-RBP) pseudogene)
match: CDS: Em:AF000364 Em:AF037448 Em:AF093821
match: ESTs: Em:AA411147 Em:W38408 Em:A1765128 Em:AV147169
Em:AA411148 Em:AA307374 Em:T89820 Em:AA677426 Em:T89909
Em:A1682619 Em:A1472757 Em:A1548569 Em:AV134944
Em:AA912195 Em:AA094744 Em:A1084864 Em:A34275 Em:AA03140
Em:A1289939 Em:D76738 Em:AA16046 Em:T39242 Em:AA407580
Em:AA561471 Em:AA647246 Em:A115211 Em:AA556549
Em:A1581394 Em:H32663 Em:A1566089 Em:AA207215 Em:AA216046
Em:T35450 Em:AA37033 Em:AA65413 Em:W9501 Em:M61969
Em:AA164877 Em:A1625087 Em:AA164209 Em:AA524815
Em:AA563205 Em:AA066208 Em:AA642158 Em:AA524815
Em:AA544174 Em:AA18195 Em:AA664286 Em:A164876
Em:A1267372 Em:AA130302 Em:AA145001 Em:AA662366
Em:AA130625 Em:AA222633 Em:AA242157 Em:AA130793
Em:AA654252 Em:AA306863 Em:AA186327 Em:AA153225
```

```
Em:AA418952 Em:T89722 Em:AA629942 Em:AA95153 Em:AA594467
Em:AV141413 Em:A1699082 Em:A1263768 Em:AA124310
Em:A1072424 Em:AAJ13602
match: Proteins: Tr:O23093 Tr:O24188 Tr:O64985 Tr:O01671
Sw:P22626 Sw:P38159 Tr:O88569 Tr:O60506 Tr:O88991
Tr:O22385 Tr:O76494 Tr:O90602 Tr:O17309 Tr:O61474
Tr:O22390 Wp:CE02193 Tr:O97560 Sw:P48810 Tr:O43347
Tr:O24668 Wp:CE03230 Tr:O76181 Tr:O41518 Tr:O43380
Tr:O17310 Sw:Q99020 Sw:P51968 Tr:Q91579 Wp:CE17059
Sw:P49314 Sw:Q13247 Tr:O04240 Tr:Q15717 Tr:O61954
Sw:Q08473 Sw:P39697 Tr:O23288 Tr:Q43349 Tr:O81987
Sw:P51991 Wp:CE00983 Tr:O08937 Tr:P90699 Tr:O15330
Sw:O01560 Tr:O24409 Tr:P70372 Tr:O61374 Tr:Q39953
Tr:P70807 Wp:CE02197 Tr:O14979"
/pseudo
/codon_start=1
/evidence=not_experimental
7549. .8004
/gene="dJ775C13.1"
/note="match: GSS: Em:AQ434110"
7614. .7732
/gene="dJ775C13.1"
/note="match: GSS: Em:AQ536515"
complement(8069. .8562)
/note="match: STS: Em:T89722"
complement(8476. .9063)
/note="match: GSS: Em:AQ008463"
8721. .8816
/note="L23 copies 3 mer att 65% conserved"
9341. .10347
/note="L1PA16 repeat: matches 5128. .6150 of consensus"
10381. .10517
/note="L1M4 repeat: matches 2211. .2348 of consensus"
11105. .11201
/note="HAL1 repeat: matches 675. .775 of consensus"
11261. .11455
/note="L2 repeat: matches 645. .869 of consensus"
11456. .11765
/note="L1M4 repeat: matches 1. .307 of consensus"
11766. .11863
/note="L2 repeat: matches 555. .645 of consensus"
12356. .12519
/note="L2 repeat: matches 298. .464 of consensus"
12520. .12615
/note="L1HS repeat: matches 6051. .6146 of consensus"
13257. .13396
/note="L70 copies 2 mer ga 69% conserved"
13442. .13893
/note="MERS4B repeat: matches 1. .498 of consensus"
13894. .13931
/note="L19 copies 2 mer ac 89% conserved"
14475. .14910
/note="L1M4 repeat: matches 5707. .6148 of consensus"
14960. .15480
/note="L1M4 repeat: matches 4030. .4551 of consensus"
15496. .16282
/note="L1PA4 repeat: matches 5361. .6146 of consensus"
16286. .16427
/note="L1PA4 repeat: matches 5213. .5358 of consensus"
16428. .16958
/note="match: GSS: Em:AQ427784"
17470. .18240
/note="L1PB1 repeat: matches 5277. .6151 of consensus"
18383. .18706
/note="AluY repeat: matches 1. .309 of consensus"
18793. .18837
/note="L15 copies 3 mer att 75% conserved"
19213. .19509
/note="AluY repeat: matches 3. .298 of consensus"
complement(join(19771. .19905),51843. .51912))
/note="match: GSS: Em:AQ377847"
22426. .22613
/note="MERS3 repeat: matches 4. .188 of consensus"
```

```

repeat_region      2385..23561
                    /note="MER5B repeat: matches 1..178 of consensus"
repeat_region      23821..23871
                    /note="L2 repeat: matches 2445..2496 of consensus"
repeat_region      23899..24082
                    /note="MER5A repeat: matches 1..181 of consensus"
repeat_region      24363..24676
                    /note="AluSk repeat: matches 3..312 of consensus"
repeat_region      25302..25481
                    /note="MLT1g repeat: matches 9..187 of consensus"
repeat_region      25486..25537
                    /note="26 copies 2 mer ct 86% conserved"
repeat_region      25564..25591
                    /note="MLT1g repeat: matches 243..270 of consensus"
repeat_region      26477..26762
                    /note="143 copies 2 mer tc 56% conserved"
repeat_region      26510..26718
                    /note="19 copies 11 mer 66% conserved"
repeat_region      26836..27041

Query Match      85.7%; Score 1442.8; DB 9; Length 103819;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 87; Indels 13; Gaps 8;

Qy 1 ATGGCTACAGAACATGTTATGGAATGTACTGGAAGAGCCCATGATGACTTACTTCAGCA 60
    |||||
Db 7351 ATGGCTACAGAACATGTTATGGAATGTACTGGAAGAGCCCATGATGACTTACTTCAGCA 7410

Qy 61 GTTATCCATTGAGAAATTTTTCAGACATTTGCTTGTGTTTACACAGAAAGTTGCT 120
    |||||
Db 7411 GTTATCCATTGAGAAATTTTTCAGACATTTGCTTGTGTTTATCACAGAAAGTTGCT 7470

Qy 121 GAAAACTAGTGAATTTAGTTGTCAGAGGCTAGTGCACATGATGATTTAGTGAAGA 180
    |||||
Db 7471 GAAAACTAGTGAATTTAGTTGTCAGAGGCTAGTGCACATGATGATTTAGTGAAGA 7530

Qy 181 GCTATCGAAGCTTTAAAGAGTTCAATGAGAGCGGCGATTTGGAGTCTTCAAGTTT 240
    |||||
Db 7531 GCTATCGAAGCTTTAAAGAGTTCAATGAGAGCGGCGATTTGGAGTCTTCAAGTTT 7589

Qy 241 AAGAGAGTGTATCTCTCATGTTTCAAGACAAAGTGCCTTTTATGTAAGTCAATGAG 300
    |||||
Db 7590 AAGAGAGTGTATCTCTCATGTTTCAAGACAAAGTGCCTTTTATGTAAGTCAATGAG 7649

Qy 301 ACTTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360
    |||||
Db 7650 CTTATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 7709

Qy 361 GAGGCAAGATTTAAGGAGCTTTTGAAGAGAGAGGCTACACATTTGATGATGACTAGAT 420
    |||||
Db 7710 GAGGCAAGATTTAAGGAGCTTTTGAAGAGAGAGGCTACACATTTGATGATGACTAGAT 7769

Qy 421 CAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
    |||||
Db 7770 CAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 7829

Qy 481 GGCAGCTGAGATTTTGTGGGAGAGATCCCCAGAGATCTGTTTGAAGATGAGCTTGTCCA 540
    |||||
Db 7830 GGCAGCTGAGATTTTGTGGGAGAGATCCCCAGAGATCTGTTTGAAGATGAGCTTGTCCA 7889

Qy 541 TTATTTGAGAGAGCTGAGAGCTATATGAGATCTTGGTTTATGATGAGATCCGCTACAGT 600
    |||||
Db 7890 TTATTTGAGAGAGCTGAGAGCTATATGAGATCTTGGTTTATGATGAGATCCGCTACAGT 7949

Qy 601 CTCAACAGAGGTTATGAGCTTTTGTCACTTTTGTACAAAGAGAGAGAGAGAGAGCTGTT 660
    |||||
Db 7950 CTCAATAGAGGTTATGAGCTTTTGTCACTTTTGTACAAAGAGAGAGAGAGAGCTGTT 8007

Qy 661 AAACGTATATATATATGAGAAATTTGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGT 720
    |||||
Db 8008 AAACGTATATATATATGAGAAATTTGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGT 8067

Qy 721 GCCAACATAGGCTTTTGTGGGCTGATCTCTTAAGAG-TAAACCAAGAGAGAGATTTCT 779

```

```

Db 8068 GCCAACATAGGCTTTTGTGGGCTGATCTCTTAAGAGTTAAACCAAGAGAGAGATTTCT 8127
    |||||
Qy 780 TGAAGATTTAGC-AAAGTACAGAGGGTCTACAGATGTCATTTTATACCAACCACTTG 838
    |||||
Db 8128 TGAAGATTTAGC-AAAGTACAGAGGGTCTTACAGAGCTCATTTTATACCAACCAACCG 8187
    |||||
Qy 839 ATGACAAAGAAAAAACAAGAGGCTTTTGTCTTTCTTGAATGAAAGATCCAAACAGCTG 898
    |||||
Db 8188 ATGACAAAGAAAAAACA-AGGCTTTTGTCTTTCTTGAATGAAAGATCCAAACAGCTG 8245
    |||||
Qy 899 CCCAGCAAGCTAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
    |||||
Db 8246 CCCAGCAAGCTAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8305
    |||||
Qy 959 TTGATGGGCTGATCCTATTTGAAGATCTGATCTGATGATGATGATGATGATGATGATGAT 1018
    |||||
Db 8306 TTGATGGGCTGATCCTATTTGAAGATCTGATCTGATGATGATGATGATGATGATGATGAT 8365
    |||||
Qy 1019 TGTTTTGAAGCAACCTTGCCACACGGTTACAGAGAAATTTTGAAGATGATGATGATGATG 1078
    |||||
Db 8366 TGTTTTGAAGCAACCTTGCCACACGGTTACAGAGAAATTTTGAAGATGATGATGATGATG 8425
    |||||
Qy 1079 AGTTTGGAAATCTGAAACGAGTAAAGAGCTAAAGATTTATGCTTTCATTCATTTGATG 1138
    |||||
Db 8426 AGTTTGGAAATCTGAAATGAGTAAAGAGCTAAAGATTTATGCTTTCATTCATTTGATG 8485
    |||||
Qy 1139 AGAGAGATGCTGTCTCAAGGCTATGAGAGAAATGAAATGATGATGATGATGATGATG 1198
    |||||
Db 8486 AGAGAGATGCTGTCTCAAGGCTATGAGAGAAATGAAATGATGATGATGATGATGATG 8545
    |||||
Qy 1199 ATATGAAATTTTGTGTTTCTTAAGCCACAGATCAGAGAGAGAGAGAGAGAGAGAGAG 1258
    |||||
Db 8546 ATATGAAATTTTGTGTTTCTTAAGCCACAGATCAGAGAGAGAGAGAGAGAGAGAGAG 8603
    |||||
Qy 1259 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1318
    |||||
Db 8604 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8663
    |||||
Qy 1319 CTCCTCCCAAGAGAGTCAAGGCTGAGAGAGTGAAGAGTGAAGATCTCCAGATT 1378
    |||||
Db 8664 CTCCTCCCAAGAGAGTCAAGGCTGAGAGAGTGAAGAGTGAAGATCTCCAGATT 8723
    |||||
Qy 1379 ATATGATACAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 1438
    |||||
Db 8724 ATATGATACAGAGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTT 8783
    |||||
Qy 1439 GATATGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1498
    |||||
Db 8784 GATATGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 8842
    |||||
Qy 1499 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558
    |||||
Db 8843 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8902
    |||||
Qy 1559 GAGCCGCTATTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1618
    |||||
Db 8903 GAGCCGCTATTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8962
    |||||
Qy 1619 GAGGTGCCCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
    |||||
Db 8963 GAGGTGCCCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9022
    |||||
Qy 1676 TGTTCACA 1683
    |||||
Db 9023 TGTTCACA 9030
    |||||

RESULT 10
AL390737 212659 bp DNA linear HTG 21-MAR-2001
LOCUS Homo sapiens chromosome 13 clone RP11-190P17, *** SEQUENCING IN
DEFINITION PROGRESS ***; 16 unordered pieces.

```

ACCESSION AL390737
VERSION AL390737.9 GI:13398796
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 212659)
REFERENCE
AUTHORS Burton, J.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Mar 21, 2001 this sequence version replaced gi:13396612.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BA190P17
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing Vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 203924 bases at least Q40
Consensus quality: 207400 bases at least Q30
Consensus quality: 209306 bases at least Q20
Insert size: 21159; sum-of-contigs
Quality coverage: 7.54x in Q20 bases; sum-of-contigs
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 48415: contig of 48415 bp in length
48416 48515: gap of 100 bp
48516 82643: contig of 34128 bp in length
82644 82743: gap of 100 bp
82744 89101: contig of 6358 bp in length
89102 89201: gap of 100 bp
89202 127367: contig of 38166 bp in length
127368 127467: gap of 100 bp
127468 169067: contig of 41600 bp in length
169068 169167: gap of 100 bp
169168 171259: contig of 2092 bp in length
171260 173159: gap of 100 bp
173160 173612: contig of 2253 bp in length
173613 173712: gap of 100 bp
173713 187445: contig of 13733 bp in length
187446 187545: gap of 100 bp
187546 191084: contig of 3539 bp in length
191085 191184: gap of 100 bp
191185 194709: contig of 3525 bp in length
194710 194809: gap of 100 bp
194810 197824: contig of 3015 bp in length
197825 197924: gap of 100 bp
197925 200623: contig of 2699 bp in length
200624 200723: gap of 100 bp
200724 203433: contig of 2710 bp in length
203434 203533: gap of 100 bp
203534 206523: contig of 2990 bp in length
206524 206623: gap of 100 bp
206624 208875: contig of 2252 bp in length
208876 212659: gap of 100 bp
212659 212659: contig of 3684 bp in length.
Location/Qualifiers
1. 212659
/organism="Homo sapiens"

/db xref="taxon:9606"
/chromosome="13"
/clone="RP11-190P17"
/clone_1b="RP11-11.1"
1. 48415
/note="assembly fragment:03097
fragment chain:1"
misc_feature
48516.82643
/note="assembly fragment:01394
fragment chain:1"
misc_feature
82744.89101
/note="assembly fragment:01816
fragment chain:1"
misc_feature
89202.127367
/note="assembly fragment:03032
fragment chain:2"
misc_feature
127468.169067
/note="assembly fragment:00218
fragment chain:2"
misc_feature
169168.171259
/note="assembly fragment:00015"
misc_feature
171360.173612
/note="assembly fragment:00477"
misc_feature
173713.187445
/note="assembly fragment:00780"
misc_feature
187546.191084
/note="assembly fragment:01634"
misc_feature
191185.194709
/note="assembly fragment:01673"
misc_feature
194810.197824
/note="assembly fragment:02018"
misc_feature
197925.200623
/note="assembly fragment:02519"
misc_feature
200724.203433
/note="assembly fragment:02839"
misc_feature
203534.206523
/note="assembly fragment:02989"
misc_feature
206624.208875
/note="assembly fragment:03934"
misc_feature
208976.212659
/note="assembly fragment:04063"
BASE COUNT 61246 a 41295 c 42666 g 65947 t 1505 others
ORIGIN
Query Match 85.7%; Score 1442.8; DB 2; Length 212659;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 87; Indels 13; Gaps 8;
QY 1 ATGGCTACGACAAATGTTAATGAAATGTTACGAGCCCATGATCTTACGCA 60
DB 83107 ATGGCTACGACAAATGTTAATGAAATGTTACGAGCCCATGATCTTACGCA 83166
QY 61 GTATCATTCAGAAATTTTACGACATTCGATGCTGTTTACGACAAAGTTGCT 120
DB 83167 GTATCATTCAGAAATTTTACGACATTCGATGCTGTTTACGACAAAGTTGCT 83226
QY 121 GAAAACTAGATGAATTTACGTTGCAAGGCTAGTTGCACATAGTATTAGTGAAGA 180
DB 83227 GAAAACTAGATGAATTTACGTTGCAAGGCTAGTTGCACATAGTATTAGTGAAGA 83286
QY 181 GCTATGGAAGCTTTAAAGAGTTCAATGAAGCGGCAATGGCAGTCTTCAACAGTTT 240
DB 83287 GCTATGGAAGCTTTAAAGAGTTCAATGAAGCGGCAATGGCAGTCTTCAACAG-TT 83345
QY 241 AAAGCAGTATCTCTCATGTTGAGAACAAAGGCTTTTATGTGAGATCAAGG 300
DB 83346 AAAGCAGTATCTCTCTCATGTTGAGAACAAAGGCTTTTATGTGAGATCAAGG 83405
QY 301 ACTTACAGGACAGAGAAAAACAGGGACCAAGTAGCAGACTTATGAAGCAAGAT 360
DB 83406 CTTATATGACAGAGAAAAACAGGGACCAAGTAGCAGACTTATGAAGCAAGAT 83465
QY 361 GAGGCAAAATTAAGGACCTTTGAAAGAACAGGCTACACACTTGATGTGACTACAGGT 420

Db 83466 GAGGCAAAATTTAAGCACTCTTGGAAGAGAGAGCTTCACTGTAATGTACCACTGGA 83525
 Qy 421 CAGAGAGATATGAGAGACCACTTCAGATTCCTTTATTCAGGTCAAGCCTCTTCTGT 480
 Db 83526 CAGAGAGATATGAGAGACCACTTCAGATTCCTTTATTCAGGTCAAGCCTCTTCTGT 83585
 Qy 481 GGCCTGAGATATTTGGGGGAGAGATCCCGAGATCTGTTGAGAGAGACTTGTCCA 540
 Db 83586 GGCCTGAGATATTTGGGGGAGAGATCCCGAGATCTGTTGAGAGAGACTTGTCCA 83645
 Qy 541 TTATTTGAGAGAGCTGAGCTATATGAGATCTGTTATATGATGATCCGCTCACTGCT 600
 Db 83646 TTATTTGAGAGAGCTGAGCTATATGAGATCTGTTATATGATGATCCGCTCACTGCT 83705
 Qy 601 CTCAACAGAGGTATGCTGTTGTCACTTTTGTACAAAGAGAGAGAGAGAGAGCTGTT 660
 Db 83706 CTCAATGAGAGGTATGCTGTTGTCACTTTTGTACAAAGAGAGAGAGAGAGCTGTT 83763
 Qy 661 AAACGTAT 720
 Db 83764 AAACGTAT 83823
 Qy 721 GCCAACAATAGGCTTTTGTGAGCTGATTCCTTAAG- TAAACCAAGAGAGAGATCTT 779
 Db 83824 GCCAACAATAGGCTTTTGTGAGCTGATTCCTTAAGATTAAGAGAGAGAGAGATCTT 83883
 Qy 780 TGAGAAATTTAGC- AAGTGAAGAGAGGTCTCAAGATGCTATTTATACCAACCACTG 838
 Db 83884 TGAAGAAATTTGCAAAAGATACAGAGGCTCTTACAGAGCTATTTATACCAACCACTG 83943
 Qy 839 ATGCAAGAAAAAAGAGAGGCTTTGCTTCTGTAATATGAGATCAAAACAGCTG 898
 Db 83944 ATGCAAGAAAAAAGAGAGGCTTTGCTTCTGTAATATGAGATCAAAACAGCTG 84001
 Qy 899 CCCAGGCAAGAGCTGAGCTTATGAGTGTAAACCAAGCTGAGGAGATTTGGAATG 958
 Db 84002 CCCAGGCAAGAGCTGAGCTTATGAGTGTAAAGTCAAGCTCTGAGGAGATTTGGAATG 84061
 Qy 959 TTGAGTGGGCTGATCTTATGAGATCTGATCTGAGAGTATGAGCAAGGTAAGTGC 1018
 Db 84062 TTGAGTGGGCTGATCTTATGAGATCTGATCTGAGAGTATGAGCAAGGTAAGTGC 84121
 Qy 1019 TGTGTTGAGAGAGCTGAGCTTATGAGTGTAAACCAAGCTGAGGAGATTTGGAATG 1078
 Db 84122 TGTGTTGAGAGAGCTGAGCTTATGAGTGTAAAGTCAAGCTCTGAGGAGATTTGGAATG 84181
 Qy 1079 AGTTGGGAACTGGAACGAGTAAAGAGCTAAAGATTAAGCTTCAATTCATTTGATG 1138
 Db 84182 AGTTGGGAACTGGAACGAGTAAAGAGCTTAAAGATTAAGCTTCAATTCATTTGATG 84241
 Qy 1139 AGAGAGATGCTGCTGCTCAAGGCTATGAGAGAGATGATGATGATGAGAGAGAG 1198
 Db 84242 AGTGAAGTGTGCTGCTCAAGGCTATGAGAGAGATGATGATGATGAGAGAGAG 84301
 Qy 1199 ATATTGAAATGTTTGTCTAAGCAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 1258
 Db 84302 ATATTGAAATGTTTGTCTAAGCAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 84359
 Qy 1259 GGCAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
 Db 84360 GGCAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84419
 Qy 1319 CTCCTCCCAAGAGAGTGAAGGCTGAGAGAGTGAAGGCTGATGATGATGATGATGAT 1378
 Db 84420 CTCCTCCCAAGAGAGTGAAGGCTGAGAGAGTGAAGGCTGATGATGATGATGATGAT 84479
 Qy 1379 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 Db 84480 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84539
 Qy 1439 GATATGAGATCCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1498

Db 84540 GATATGAAGATCCTACTATGATGATGATGATGATGATGATGATGATGATGATGAT 84598
 Qy 1499 GTAGAGAGAGAGAGAGTGTGCTGCTTCATCCAGAGGTGTGGGGCTGCTCTCCCTGTGTA 1558
 Db 84599 GTAGAGAGAGAGAGAGTGTGCTGCTTCATCCAGAGGTGTGGGGCTGCTCTCCCTGTGTA 84658
 Qy 1559 GAGCCGGTATATCAAGAGAGAGAGGCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 1618
 Db 84659 GAGCCGGTATATCAAGAGAGAGAGGCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 84718
 Qy 1619 GAGGTGCCAACAACAAAG 1675
 Db 84719 GAGGTGCCAACAACAAAG 84778
 Qy 1676 TGTTCACA 1683
 Db 84779 TGTTCACA 84786
 RESULT 11
 BC032643
 LOCUS
 DEFINITION
 Homo sapiens, similar to NSI-associated protein 1, clone MGC:45213
 IMAGE:5495201, mRNA, complete cds.
 ACCESSION
 BC032643
 VERSION
 BC032643.1 GI:21619167
 KEYWORDS
 MGC.
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3319)
 REFERENCE
 AUTHORS
 Strausberg, R.
 TITLE
 Direct Submision
 JOURNAL
 Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT
 Contact: MGC Help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Ahlert, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantirp, S., Thomas, P.J., Touchman, J.W.,
 Tsurgan, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clome distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINT at: <http://image.llnl.gov>
 Series: IRK Plate: 69 Row: f Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5455805.
 Location/Qualifiers
 1. 3319
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:45213 IMAGE:5495201"
 /tissue_type="Eye, retinoblastoma"
 /clone_lib="NIH MGC_67"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 98. 1681
 CDS

LOCUS AF441128 1899 bp mRNA Linear ROD 04-JAN-2002
 DEFINITION Mus musculus CD-1 heterogeneous nuclear ribonucleoprotein R mRNA,
 complete cds.
 ACCESSION AF441128
 VERSION AF441128.1 GI:17066598
 KEYWORDS Mus musculus.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1899)
 Rosboll, W., Kroning, A.K., Ohndorf, U.M., Steegborn, C., Jablonka, S.
 and Sendtner, M.
 TITLE Specific interaction of Smn, the spinal muscular atrophy
 determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role
 for Smn in RNA processing in motor axons?
 Hum. Mol. Genet. 11 (1), 93-105 (2002)
 JOURNAL Hum. Mol. Genet. 11 (1), 93-105 (2002)
 MEDLINE 21634625
 PUBMED 11773003
 REFERENCE 2 (bases 1 to 1899)
 Rosboll, W., Kroning, A.-K. and Sendtner, M.
 TITLE Direct Submission
 Submitted (30-OCT-2001) Inst. f. Clinical Neurobiology, Muenzberg
 University, Josef-Schneider Str. 11, Muenzberg 97080, Germany
 FEATURES
 source
 1. 1899
 /organism="Mus musculus"
 /strain="CD-1"
 /db_xref="taxon:10090"
 1. 1899
 /note="hnRNP-R"
 /codon_start=1
 /product="heterogeneous nuclear ribonucleoprotein R"
 /protein_id="AA15332.1"
 /db_xref="GI:17066598"
 /translation="MANQVNGNAVOLKEESEPMDTSVTHTEHYKTLIEAGLPQKVAE
 RLDEIFQTVGLVADLDERAIDALREFNEBGLSVLOFPESDLSHYONSAFLGVN
 KTYRREKSGKSGKQESTKPEAKIKALIRTYGTLVTTGQRYGGPPDSVYSGV
 PGCTEYFNKIPKIPRLDYEDDELVPIPERAGPIWDLRLMDPLSGNRYATITTCGKA
 AOEAVKLCDSLEIRPKALGVCISVANRRLEVGSIPNKKENTLEFSVTBLVDV
 ILVHNPDKKRNKFCFLYEYEDHNSAAQARRRLMSGVKNVNVYEMADPVEEDP
 EVNAKVLFLFVRNLATVTBEILIEKSEFSEKSLKSDYAVFVHEDEGAALVAMD
 EMNGKEIEGEIEIVLAPDKPKRKERQAQASRSTAYEDYVHPPPRPPRGRG
 RGGKGYGYPDYEDYDDYDGYDHYDGGEDPYGYDGYDGYAVRGGGGRGR
 GAPPPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRGAPRG
 NUGKRAKADGTNDPSKRRKRTNMQNMGSOP1AQDPLQGGSDYSNGYNNDDQEFYQ
 DTYGQQWK"
 BASE COUNT 581 a 362 c 548 g 408 t
 ORIGIN
 Query Match 53.2%; Score 896; DB 10; Length 1899;
 Best Local Similarity 73.6%; Pred. No. 1,2e-193;
 Matches 1184; Conservative 0; Mismatches 415; Indels 9; Gaps 3;

351 AGAACGATGAGCAAGATTAAGCACTTTGAAAAGAACGCTACACTTGAAT 410
 Db AGGGCCGATGAGCAAGATTAAGCACTTTGAAAAGAACGCTACACTTGAAT 419
 Qy GACTACAGATCAAGAGATATGAGAGACCACTCCAGATTCGTTTATTCAGGTGCA 470
 Db GACCAAGGACAGAGAGATATGAGAGACCACTCCAGATTCGTTTATTCAGGTGCA 479
 Qy GCCTTCGTTGGCACTGAGATATTTTGGGAAAGATCCCAAGATCTGTTGAGATGA 530
 Db GCCTGAGATGGGACAGAGGCTTCCTGAGTAATATCCAAAGACCTTGATGAGATGA 539
 Qy GCTTGTCCATTAATTTGAGAAAGCTGACCTTAATGAGATCTTCCTTAATGATATCC 590
 Db ATTATGATCCCTTTTGAAGAGGCTGCTATCTGGGAATCTAGGCTCATGATGATCC 599
 Qy GCTCACTGATCTCAACAGAGTTATGCGTTGTCACTTTTGTACAAAGAACAGACACA 650
 Db ACTCTCTGTCAGAACAGAGGCTATGATTCATCACTTCTGTGAAAAGAACGTCGCA 659
 Qy AGAGGCTGTTAACTGTATATATATCAATGAATTCGTTCCGGAAGACATTTGCTGTG 710
 Db GAAAGCAGTTAAGCTGTGTGACAGCTATGAGATTCGCTGTAAACACTGGAGAGTGTG 719
 Qy CATCTCAGTTGCCAACAATAGCTTTTGTGGGCTCGATTCCTTAAGATTAACCAAGA 770
 Db CATTTCTGTGCAACAAAGACTTTTGTGTGATCCATTCACAAAGATTAACCAAGA 779
 Qy GCAGATCTTGAAGATTTTGAAGATGACAGAGGCTCAGAGATGCTATTTATATCA 830
 Db GAACATCTTGGAAGATTCAGTAAAGTCACAGAGGTTTGTGACGTTATTTCTATCA 839
 Qy CCAACTGATGACAAAGAAAAAACAAGGCTTTTGTCTTTCTTGAATATGAATCAAA 890
 Db TCAACCCGATGACAAAGAAAAAAGAATCGGGGCTTCGCTTCCTTGAATATGAGATCAAA 899
 Qy AAGAGCTGCCCAAGCAAGATGAGCTTAATGAATGCTTAAGTCAAACTCGGGAAATGT 950
 Db GTCCAGCAGCAACAAAGCAAGCGCGCTGATGAGTGAAGAAAGTCAAACTGCGGAAATGT 959
 Qy TGGAACTGTTGAGTGGGCTGATCCTATTAAGAATCTGATCCTGAAGTTATGCAAAAGT 1010
 Db GGTTCAGTAGAGTGGGCTGATCCTGATGAGAGCCGATCCAGAAATCATGGCAAAAGT 1019
 Qy AAAAGTCTGTTGTACGCAACCTTGCACACAGTAAACAGAAATTTTGAAGAAATGTC 1070
 Db AAAAGTCTGTTGTACGCAACCTTGCACACAGTAAACAGAAATTTTGAAGAAATGTC 1079
 Qy ATTATGACATTTGGGAAACTGGAACAGATTAAGAACTTAAGATTAATGCTTCAATCA 1130
 Db ATTTTCTGAGTTTGGAAACTGGAACAGATTAAGAACTTAAGATTAATGCTTCAATCA 1139
 Qy TTTTGTATGAGAGATGATGCTGCTCAAGGCTATGAGAAATGATGATTAAGATCTTGA 1190
 Db TTTTGAAGACAGGAGAGCTGCTGTTAAGCTATGAGAAATGATGATTAAGAAATGAG 1199
 Qy GGGAGAAATATTTGAAATTTGTTTGTCTAAGCAACAGATCAGAAAGAGAAAGAAA 1250
 Db AGGGAGAAATATTTGAAATTTGTTTGTCTAAGCAACAGATCAGAAAGAGAAAGAAA 1259
 Qy AGCTCAGAGCAGCAGCAGCAAGAAATCAAAATGATGATTAATCACTATATATGCTCAC 1310
 Db AGCTGCTAGACAGGCTCCTCAGAGCACTGCTGATGAAATATTTATTAACCTCTCTCC 1319
 Qy TCATATGCTCTCCCAACAAGAGTGAAGGCTGAGAGTGAAGAGTGGCTATGATATCC 1370
 Db CCGAATGCTCACTCAATGAGAGGCTGATGAGAGGCTGATGAGAGTGGCTATGATATCC 1379
 Qy TCCAGATTAATATGATATGAGAG---ATTATTAATATTAATATGATATGATATCAATA 1427
 Db TCCAGATTAATATGATATGAGAGCTATGATATGATATTAATGATATGATATCAATA 1439

QY 1428 CTATCGTGTGATGATGAGATCCATCTATGTTATGAGAT---TTTCAAGTTGGAGC 1484
 DB 1440 CTATCGTGTGAGGCTATGATGATGCTTACTACGCTATGATGATGCTATGATGATGAGG 1499
 QY 1485 TAGAGGAGGGGTGTGATGAGGAGGAGGGGTGCTGCTTCATTCAGAGGTGCTGGGCTGC 1544
 DB 1500 AAGAGGAGGAGGAGGAGGGGGGGGGGTGCTCCACACACAGAGGGGGGGGGAGCACC 1559
 QY 1545 TCCCTCCCGGAGTGTAGAGCCGGTTATTCAGAGAGAGAGCCCT---GATCAGCAAGAGG 1601
 DB 1560 ACCTTCAGAGGAGTGTAGAGCTGCTATTTTCAAGAGGGGGGACCTTTGGACACCAAGAGG 1619
 QY 1602 CGTTCGGGTGTGAGGAGAGGAGTCCCAACAACAAGAGGCGCGGGG 1649
 DB 1620 CTCAGAGGGGTGTGAGGAGGGGGTCTCTGCAACAAGAGAGGCGCTGG 1667
 RESULT 15
 AC118904/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-234J3, *** SEQUENCING IN PROGRESS
 *** 81 unordered pieces.
 AC118904 151297 bp DNA linear HTG 23-JUL-2002
 AC118904
 AC118904.2 GI:21903156
 VERSION HTG; HTGS PHAS1.
 KEYWORDS Rattus norvegicus.
 SOURCE Rattus norvegicus.
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 151297)
 Muzny,D.M., Adams,C., Adio-oduola,B., Ali-osman,F.R., Allen,C., Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbieri,J., Benton,V., Bimberg,K., Blankenburg,K., Bonlin,D., Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garra,N., Gill,R., Correil,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louleaged,H., Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabac,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,S., Ogul,M., Okunon,G., Nguyen,N., Nickerson,B., Nwokenkwo,S., Ogul,M., Okunon,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primas,E., Pul,L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Roife,M., Ruiz,S., Severy,G., Sodergren,B., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansun,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umani,K., Vazquez,L., Vera,V., Villalon,D., Vinson,R., Wang,C., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.
 TITLE
 JOURNAL
 Unpublished

REFERENCE 2 (bases 1 to 151297)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 151297)
 REFERENCE
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 18, 2002 this sequence version replaced gi:20258378.
 COMMENT
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GWIS
 Center clone name: CH230-234J3
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 82728 bases at least Q40
 Consensus quality: 89631 bases at least Q30
 Consensus quality: 95421 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 81 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1
 1029: contig of 1029 bp in length
 1030
 1129: gap of unknown length
 1130
 2205: contig of 1076 bp in length
 2206
 2305: gap of unknown length
 2306
 3335: contig of 1030 bp in length
 3336
 3435: gap of unknown length
 3436
 4450: contig of 1015 bp in length
 4451
 4550: gap of unknown length
 4551
 5562: contig of 1012 bp in length
 5563
 6689: gap of unknown length
 6690
 6789: gap of unknown length
 6790
 7789: gap of unknown length
 7790
 9261: contig of 1372 bp in length
 7890
 9361: gap of unknown length
 9362
 10406: contig of 1045 bp in length
 10407
 10506: gap of unknown length
 10507
 11513: contig of 1007 bp in length
 11514
 11613: gap of unknown length
 11614
 12691: contig of 1078 bp in length
 12692
 12791: gap of unknown length
 12792
 13853: contig of 1062 bp in length
 13854
 15093: gap of unknown length
 15094
 15192: gap of unknown length
 15193
 16211: contig of 1019 bp in length
 16212
 16311: gap of unknown length
 16312
 17349: contig of 1038 bp in length
 17350
 17449: gap of unknown length
 17450
 18672: contig of 1223 bp in length
 18673
 18772: gap of unknown length
 18773
 20353: contig of 1581 bp in length
 20354
 20453: gap of unknown length

20454 21476: contig of 1023 bp in length
21477 21575: gap of unknown length
21577 22610: contig of 1034 bp in length
22611 22710: gap of unknown length
22711 23713: contig of 1003 bp in length
23714 23813: gap of unknown length
23814 25236: contig of 1423 bp in length
25237 25337: gap of unknown length
25337 26357: contig of 1021 bp in length
26358 26457: gap of unknown length
26458 27597: contig of 1140 bp in length
27598 28843: contig of 1146 bp in length
28844 28944: gap of unknown length
28944 30133: contig of 1196 bp in length
30140 30239: gap of unknown length
30240 31239: contig of 1000 bp in length
31240 31339: gap of unknown length
31340 32369: contig of 1030 bp in length
32370 32459: gap of unknown length
32470 33484: contig of 1015 bp in length
33485 33584: gap of unknown length
33585 35005: contig of 1421 bp in length
35006 35105: gap of unknown length
35106 36271: contig of 1166 bp in length
36272 36371: gap of unknown length
36372 37533: contig of 1162 bp in length
37534 37633: gap of unknown length
37634 38809: contig of 1176 bp in length
38810 38909: gap of unknown length
38910 40148: contig of 1239 bp in length
40149 40248: gap of unknown length
40249 41411: contig of 1163 bp in length
41412 41512: gap of unknown length
41512 43868: contig of 2357 bp in length
43869 45611: contig of 1643 bp in length
45612 45711: gap of unknown length
45712 47108: contig of 1397 bp in length
47109 47208: gap of unknown length
47209 48273: contig of 1065 bp in length
48274 48373: gap of unknown length
48374 50006: contig of 1633 bp in length
50007 50107: gap of unknown length
50107 51371: contig of 1265 bp in length
51372 51471: gap of unknown length
51472 52752: contig of 1280 bp in length
52753 52851: gap of unknown length
52852 53943: contig of 1092 bp in length
53944 54043: gap of unknown length
54044 56187: contig of 2144 bp in length
56188 56287: gap of unknown length
56288 57740: contig of 1453 bp in length
57741 57840: gap of unknown length
57841 58957: contig of 1117 bp in length
58958 59057: gap of unknown length
59058 60657: contig of 1500 bp in length
60658 61991: contig of 1334 bp in length
61992 62092: gap of unknown length
62093 63961: contig of 1870 bp in length
63962 64061: gap of unknown length
64062 65338: contig of 1277 bp in length
65339 65438: gap of unknown length
65439 66681: contig of 1243 bp in length
66682 66781: gap of unknown length
66782 68027: contig of 1246 bp in length
68028 68127: gap of unknown length
68128 69486: contig of 1359 bp in length
69487 69586: gap of unknown length
69587 71279: contig of 1692 bp in length
71279 71379: gap of unknown length
71379 73777: contig of 2399 bp in length

Query Match 53.2%; Score 895.8; DB 2; Length 151297;
Best Local Similarity 79.3%; Pred. No. 2.1e-193;
Matches 1291; Conservative 0; Mismatches 279; Indels 59; Gaps 17;
QY 59 CAGTATCCATTCAGAAATTTTCAGACATTCCTGATGCTGTTACACAGAAAGTTG 118
DB 43868 CACTATCCCATCCACAAATTTTCAGACCT--TTGATGCTTAATATCCAAAGTTG 43811
QY 119 CTGAAAACTAGATGAATTTTACGTTGACAGGGCTAGTTGACATAGATTAGATGAA 178
DB 43810 CTGAAANCTAGATGAATTTTAG-TCGTGGCTAGTTGACATAGATTTAGATGAAA 43752
QY 179 GAGCTATCCAGCTTTAAAGAGTTCATATGAAACGCGATTTGGAGCTTCAACAG 238
DB 43751 AAGCTATCAAGCTTAAAGAGTTCATATGAAACATATGATGATGACAGCTTCAACAT 43692
QY 239 TTAAGACAGTATCTCTCATGTTTCAGAAACAAAGTCTTTTATGTCAGTCATGA 238
DB 43691 TTAAGACTGTATCTCTCTGTCG-----CCCTTTTATGTCAGTCATGA 43646
QY 299 AGACTTACAGCAG 358
DB 43645 AGACTTACAGCAG 43587
QY 359 ATGAGGAAAGATTTAAGGACCTTTTGGAAAGAGAGAGAGAGAGAGAGAGAGAG 418
DB 43586 ATGAGGAAAGATTTAAGTCACTCTGTGTAGAGAGAGAGAGAGAGAGAGAGAG 43527
QY 419 GTGAGAGAGATTTGAG 477
DB 43526 GACAGAGAGATTTAG 43467
QY 478 GTTGCACTGAGATTTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
DB 43466 GTTGCACTGAGATTTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 43407
QY 538 CCAATTTTGAAG 537
DB 43406 CCAATTTTGAAG 43347
QY 598 GGTCTCAAGAGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
DB 43346 GGTCTCAAGAGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 43290
QY 658 GTTAACTGTAT 717
DB 43289 GTTAACTGT-----ATCATGAAATCATATTTGAAAGAGAGAGAGAGAGAG 43236
QY 718 GTTGCCCAAGATAGGCTTTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
DB 43235 GTTGCCCAAGATAGG--GGGGGGGGGCTCGTTTCAAGAGAGAGAGAGAGAGAT 43178
QY 778 CTTGAGGAATTTAGCAAG 837
DB 43177 CTTGAGGAATTTAGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 43118
QY 838 GATGACAAAG 897
DB 43117 GATGACAAAG 43059
QY 898 GCCCAG 957
DB 43058 GCCCAG 42999
QY 958 GTTGAAG 1017
DB 42998 GTTGAAG 42945
QY 1018 CTTGTTTGAAG 1077
DB 42944 CTTGTTTGAAG 42885

THIS PAGE BLANK (USPTO)